

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
MESTRADO PROFISSIONAL EM ADMINISTRAÇÃO PÚBLICA

Inteligência Artificial Generativa nos processos licitatórios do setor público: um estudo de caso no TJPE sobre a engenharia de *prompt* para artefatos de licitação e seus aspectos éticos e de transparência

Liana Beatriz dos Santos Barreto de Souza

Recife, julho de 2025

Liana Beatriz dos Santos Barreto de Souza

Inteligência Artificial Generativa nos processos licitatórios do setor público: um estudo de caso no TJPE sobre a engenharia de *prompt* para artefatos de licitação e seus aspectos éticos e de transparência

Orientador: Jorge da Silva Correia Neto

Dissertação apresentada ao Curso de Mestrado Profissional em Administração Pública da Universidade Federal Rural de Pernambuco como requisito parcial para conclusão do Curso.

Recife, julho de 2025

RESUMO

A implementação da Inteligência Artificial Generativa (IAG) no Poder Judiciário brasileiro emerge como resposta aos desafios de modernização e eficiência diante do alto volume processual e da complexidade burocrática. Contudo, sua aplicação enfrenta obstáculos significativos relacionados à falta de transparência, vieses algorítmicos e o fenômeno das alucinações. Este estudo concentra-se no Tribunal de Justiça de Pernambuco (TJPE) e na modalidade de Dispensa de Licitação, formulando a seguinte questão: de que modo a engenharia de *prompts* pode ser estruturada para viabilizar o uso ético e transparente da IAG na elaboração de artefatos administrativos voltados à contratação direta por Dispensa de Licitação no âmbito do TJPE? O objetivo geral foi investigar como a engenharia de *prompts* pode ser estruturada para viabilizar o uso ético e transparente da IAG na elaboração de artefatos administrativos voltados à contratação direta por Dispensa de Licitação. A metodologia caracteriza-se como pesquisa aplicada, exploratório-descritiva, com abordagem mista e delineamento experimental. Desenvolveu-se uma arquitetura híbrida integrando cinco técnicas: *Chain-of-Thought* para estruturação do raciocínio jurídico-administrativo, *Retrieval-Augmented Generation* para consulta a bases normativas, *Self-Refine* para controle de qualidade automatizado, *Few-Shot Learning* para padronização institucional, complementadas pela configuração de Persona e temperatura baixa ($\leq$0.2). Limitações do ChatGPT conduziram à seleção do Claude (Anthropic) para experimentação. A validação empírica consistiu na geração de cinco Termos de Referência representativos, avaliados por três especialistas mediante instrumento estruturado. Os resultados demonstram eficácia significativa: *score* médio de 93,7% de conformidade legal, redução temporal de 99,92% (de 60 horas para 2 minutos e 30 segundos), e ausência completa de alucinações. A análise dimensional revelou excelência na Qualidade Técnica (99,1%) e Conformidade Institucional (99,6%), com convergência excepcional entre avaliadores (coeficiente de variação de 2,4%). As técnicas implementadas mostraram-se eficazes na mitigação de alucinações, embora a supervisão humana permaneça indispensável. Conclui-se que a pesquisa valida empiricamente a aplicabilidade da IAG para otimização de processos administrativos no judiciário, contribuindo para maior eficiência operacional e segurança jurídica. O Produto Técnico-Tecnológico resultante constitui guia metodológico para adoção segura da IAG em outras instituições públicas, alinhado à Resolução CNJ nº 615/2025, com potencial de replicabilidade nacional e impacto transformador na modernização da administração pública brasileira.

Palavras-chave: Inteligência Artificial Generativa; Engenharia de *Prompts*; Licitações Públicas; Poder Judiciário; Ética e Transparência.

ABSTRACT

The implementation of Generative Artificial Intelligence (GAI) in the Brazilian Judiciary emerges as a strategic response to modernization and efficiency challenges given the high procedural volume and bureaucratic complexity. However, its application faces significant obstacles related to lack of transparency, algorithmic biases, and the hallucination phenomenon. This study focuses on the Court of Justice of Pernambuco (TJPE) and the Tender Exemption modality, formulating the following question: how can prompt engineering be structured to enable ethical and transparent use of GAI in preparing administrative artifacts for direct contracting through Tender Exemption within TJPE? The general objective was to investigate how prompt engineering can be structured to enable ethical and transparent use of GAI in preparing administrative artifacts for direct contracting through Tender Exemption. The methodology is characterized as applied, exploratory-descriptive research with a mixed approach and experimental design. A hybrid architecture was developed integrating five techniques: Chain-of-Thought for structuring legal-administrative reasoning, Retrieval-Augmented Generation for consulting updated regulatory databases, Self-Refine for automated quality control, Few-Shot Learning for institutional standardization, complemented by Persona configuration and low temperature (≤0.2). ChatGPT limitations led to selecting Claude (Anthropic) for experimentation. Empirical validation consisted of generating five representative Terms of Reference, evaluated by three experts using a structured instrument. Results demonstrate significant effectiveness: average score of 93.7% legal compliance, 99.92% time reduction (from 60 hours to 2 minutes and 30 seconds), and complete absence of hallucinations. Dimensional analysis revealed excellence in Technical Quality (99.1%) and Institutional Compliance (99.6%), with exceptional convergence among evaluators (coefficient of variation 2.4%). Implemented techniques proved effective in mitigating hallucinations, although human supervision remains indispensable. The research empirically validates GAI applicability for optimizing administrative processes in the judiciary, contributing to greater operational efficiency and legal certainty. The resulting Technical-Technological Product constitutes a methodological guide for safe GAI adoption in other public institutions, aligned with CNJ Resolution No. 615/2025, with national replicability potential and transformative impact on Brazilian public administration modernization.

Keywords: Generative Artificial Intelligence; Prompt Engineering; Public Procurement; Judiciary; Ethics and Transparency.

Lista de Figuras

Figura 1: Objetivos geral e específicos.....5
Figura 2: Inteligência Artificial Generativa nos processos licitatórios do setor público.....6
Figura 3: Desafios de implementação do IA no judiciário.....14
Figura 4: Justificativa do projeto.....15
Figura 5: Percentual de uso de IAG.....18
Figura 6: Principais limitações, dificuldades ou desafios ao se utilizar ferramentas de IAG ..19
Figura 7: Panorama da Crise do Poder Judiciário Brasileiro em 2023: volume processual, custos operacionais, percepção pública e emergência de soluções tecnológicas baseadas em Inteligência Artificial.....26
Figura 8: Índice de produtividade por magistrado.....26
Figura 9: Índice de produtividade por servidor. 26
Figura 10: Dashboard de Indicadores de Eficiência do Poder Judiciário Brasileiro (2023): impacto do Programa Justiça 4.0 na Taxa de Congestionamento, Índice de Atendimento à Demanda e Índice de Produtividade da Magistratura.....27
Figura 11: Os quatro eixos do Programa Justiça 4.0 29
Figura 12: Percentual de servidores(as) e/ou assessores(as) que informaram ter utilizado a IAG como ferramenta de apoio no desenvolvimento das atividades profissionais.....30
Figura 13: Projetos de IA em 2024, por ramo de justiça.....32
Figura 14: Distribuição territorial dos Tribunais de Justiça segundo o porte.....35
Figura 15: Licitações homologadas no TJPE por modalidade e ano (2022–2024).....36
Figura 16: Total de projetos de IA no judiciário brasileiro entre 2022 e 2024 37
Figura 17: Subáreas de IA utilizadas.....38
Figura 18: Principais desafios no uso de IA no judiciário.....39
Figura 19: Mecanismo do Transformer.....51
Figura 20: Diagrama do mecanismo Self-attention.....52
Figura 21: Arquitetura do LLM.....53
Figura 22: Taxonomia de técnicas de engenharia de prompts.....55
Figura 23: Fenômeno das alucinações em IAG.....64
Figura 24: IAGs utilizadas no Poder Judiciário Brasileiro.....78
Figura 25: Exemplo de limitação do ChatGPT na geração completa de Termos de Referência 79
Figura 26: Design Metodológico - Desenvolvimento de Prompts para IAG em Licitações Públicas.....82
Figura 27: Arquitetura de prompt híbrido proposto 86

Lista de Quadros

Quadro 1: Timeline da Transformação Digital no Judiciário Brasileiro 7
Quadro 2:Terminologia de Segurança em Prompting 9
Quadro 3: Abordagens Conceituais da Inteligência Artificial..... 21
Quadro 4: Perfil de Transformação Digital dos Principais Tribunais de Justiça..... 35
Quadro 5: Resolução CNJ nº 332/2020 × Resolução CNJ nº 615/2025..... 40
Quadro 6: Diferenças entre modelos de inteligência artificial descritivos e generativos..... 50
Quadro 7: Taxonomia contemporânea dos prompts..... 54
Quadro 8: Técnicas de engenharia de prompt 57
Quadro 9: Comparativo entre as abordagens Role-Playing e Personalization na Engenharia de Prompts com Persona 58
Quadro 10: Efeitos da Temperatura em Modelos de Linguagem..... 60
Quadro 11: Detalhamento dos dez elementos obrigatórios do Termo de Referência (Art. 6º, XXIII, Lei nº 14.133/2021) 67

Lista de Tabelas

Tabela 1: Scores Consolidados por Avaliador e TR.....94
Tabela 2: Scores de Completude Normativa por TR e Avaliador.....96
Tabela 3: Scores de Qualidade Técnica por TR e Avaliador.....97
Tabela 4: Scores de Conformidade Institucional por TR e Avaliador.....98
Tabela 5: Análise de Variabilidade entre Avaliadores.....99
Tabela 6: Incidência de Alucinações por TR.....99
Tabela 7: Comparação Temporal Detalhada.....100

Lista de Abreviaturas e Siglas

ChatGPT - *Chat Generative Pre-Trained Transformer*

CNJ - Conselho Nacional de Justiça

CPC - Código de Processo Civil de 2015

CRFB - Constituição da República Federativa do Brasil de 1988 e emendas constitucionais posteriores

Datajud - Base Nacional de Dados do Poder Judiciário

IA - Inteligência Artificial

IAG - Inteligência Artificial Generativa

ICJBrasil - Índice de Confiança na Justiça no Brasil

LLM – *Large Language Model* (Grande Modelo de Linguagem)

ODS - Objetivos de Desenvolvimento Sustentável da Agenda 2030

PLN – Processamento de Linguagem Natural

PDPJ-Br - Plataforma Digital do Poder Judiciário Brasileiro

PNUD - Programa das Nações Unidas para o Desenvolvimento

RAG - *Retrieval-Augmented Generation* (Geração Aumentada por Recuperação)

RNAs - Redes Neurais Artificiais

TCU - Tribunal de Contas da União

TJPE - Tribunal de Justiça de Pernambuco

SUMÁRIO

1 INTRODUÇÃO INTEGRATIVA 1
1.1 CONTEXTUALIZAÇÃO 1
1.2 OBJETIVOS 4
1.2.1 Objetivo Geral 4
1.2.2 Objetivos Específicos 4
1.3 JUSTIFICATIVA 5
1.3.1 Aprimoramento da Gestão Pública e Modernização do Judiciário 11
1.3.2 Desafios Éticos e Jurídicos 12
1.3.3 Relevância Acadêmica e Social 14
1.4 ESTRUTURA DA DISSERTAÇÃO 15
2. INTELIGÊNCIA ARTIFICIAL NO JUDICIÁRIO BRASILEIRO: DESAFIOS E OPORTUNIDADES ACERCA DOS ASPECTOS ÉTICOS E DE TRANSPARÊNCIA 17
2.1. INTRODUÇÃO 17
2.2. INTELIGÊNCIA ARTIFICIAL 20
2.3. PROCESSAMENTO DE LINGUAGEM NATURAL E GRANDE MODELO DE LINGUAGEM 22
2.4. A INTELIGÊNCIA ARTIFICIAL E O JUDICIÁRIO BRASILEIRO 24
2.5. A TRANSFORMAÇÃO DIGITAL DO PODER JUDICIÁRIO 27
2.5.1. Programa Justiça 4.0 - A Revolução Judicial em Curso 28
2.5.2. Plataforma Sinapses: O Núcleo da Inteligência Artificial no Judiciário 30
2.5.3. Governança e Transparência Algorítmica 31
2.6. A EXPERIÊNCIA DO TJPE NA TRANSFORMAÇÃO DIGITAL E NO USO DA INTELIGÊNCIA ARTIFICIAL 33
2.7. DESAFIOS E PERSPECTIVAS ACERCA DA ÉTICA E DA TRANSPARÊNCIA 36
REFERÊNCIAS 41
3. INTELIGÊNCIA ARTIFICIAL GENERATIVA NAS LICITAÇÕES PÚBLICAS: REQUISITOS PARA ESTRUTURAÇÃO DOS *PROMPTS* DE DISPENSA DE LICITAÇÃO 47
3.1. INTRODUÇÃO 47
3.2. INTELIGÊNCIA ARTIFICIAL GENERATIVA 49
3.2.1. Transformer 50
3.2.2. Prompts 53
3.2.2.1. Evolução e Classificação dos *Prompts* 54
3.2.2.2. Técnicas Avançadas de Engenharia de *Prompts* 55

3.2.2.3. Parâmetros de Configuração de *Prompts* ... 57
3.2.2.4. Integração de Técnicas e Arquiteturas Híbridas ... 60
3.2.3. Alucinações ... 61
3.3. FUNDAMENTOS DA LICITAÇÃO ... 65
3.3.1. Conceito, Natureza Jurídica e Fundamentos Constitucionais ... 65
3.3.2. Lei nº 14.133/2021 - Nova Lei de Licitações ... 65
3.3.3. Modalidades de Compras Diretas ... 66
3.3.4. Termo de Referência: Artefato Central da Pesquisa ... 66
3.4. CONSIDERAÇÕES FINAIS ... 68
REFERÊNCIAS ... 70
4. INTELIGÊNCIA ARTIFICIAL GENERATIVA NAS LICITAÇÕES PÚBLICAS: UM ESTUDO DE CASO NO TRIBUNAL DE JUSTIÇA DE PERNAMBUCO ... 73
4.1. INTRODUÇÃO ... 73
4.2. METODOLOGIA DA PESQUISA ... 74
4.2.1. Classificação ... 74
4.2.2. Procedimentos Metodológicos ... 76
4.2.2.1. Primeira Etapa: Fundamentação Teórica e Normativa ... 76
4.2.2.2. Segunda Etapa: Seleção e Adaptação de Técnicas de Engenharia de *Prompts* ... 77
4.2.2.3. Terceira Etapa: Seleção de Ferramentas de IAG ... 78
4.2.2.3.1. Limitações identificadas no ChatGPT ... 79
4.2.2.3.2. Transição para o Claude ... 80
4.2.2.4. Quarta Etapa: Desenvolvimento e Validação Experimental ... 80
4.2.3. Limitações Metodológicas ... 82
4.3. DESENVOLVIMENTO DOS *PROMPTS* ... 84
4.3.1. Arquitetura Híbrida Implementada ... 85
4.3.2. Processo de Refinamento Iterativo ... 87
4.3.3. Estrutura Final dos *Prompts* Validados ... 87
4.3.4. Validação Técnica e Ajustes ... 88
4.4. EXPERIMENTAÇÃO E COLETA DE DADOS ... 89
4.4.1. Coleta e Análise de Dados ... 89
4.4.2. Instrumento de Avaliação (*checklis*t) ... 91
4.5. RESULTADOS E ANÁLISE ... 92
4.5.1. Estrutura de Avaliação e Cálculo de *Scores* ... 92
4.5.1.1. Sistema de Pontuação por Dimensão ... 92
4.5.1.2. Justificativa dos Pesos Atribuídos ... 93
4.5.2. Análise de Resultados por Avaliador ... 93
4.5.2.1. *Performance* por Avaliador ... 94

4.5.2.2. Cálculo Detalhado dos *Scores* 94
4.5.3. Análise por Dimensão de Avaliação 95
4.5.3.1. Completude Normativa 95
4.5.3.2. Qualidade Técnica 96
4.5.3.3. Conformidade Institucional 97
4.5.4. Análise de Concordância entre Avaliadores 98
4.5.5. Controle de Alucinações 99
4.5.6. Eficiência Temporal 100
4.5.7. Síntese dos Resultados 100
4.6. DISCUSSÃO DOS ACHADOS 101
4.6.1. Avaliação da Qualidade dos Documentos Gerados 101
4.6.2. Eficiência Temporal e Operacional 102
4.6.3. Validação da Confiabilidade através do Coeficiente de Variação 103
4.6.3.1. Conceituação do Coeficiente de Variação 103
4.6.3.2. Validação dos Resultados 104
4.6.4. Análise Dimensional da Performance: *Insights* Específicos por Categoria 104
4.6.4.1. Completude Normativa (85,3%): Maior Variabilidade e Oportunidades de Refinamento 104
4.6.4.2. Qualidade Técnica (99,1%): Excelência na Estruturação do Raciocínio 104
4.6.4.3. Conformidade Institucional (99,6%): Sucesso da Configuração de Persona 105
4.6.4.4. Implicações Metodológicas das Variações Dimensionais 105
4.6.5. Mitigação de Alucinações e Controle de Qualidade 105
4.6.6. Transição Tecnológica: ChatGPT para Claude 106
4.6.7. Conformidade com Marcos Normativos e Transparência 106
4.6.8. Validação da Arquitetura Híbrida de *Prompts* 107
4.6.9. Aceitação e Usabilidade Institucional 107
4.6.10. Limitações e Implicações para Pesquisas Futuras 107
4.6.11. Contribuições para a Governança Algorítmica 108
4.6.12. Implicações para a Transformação Digital do Judiciário 108
4.7. CONSIDERAÇÕES FINAIS 109
5. SÍNTESE DO PRODUTO TÉCNICO-TECNOLÓGICO . **Erro! Indicador não definido.**
5.1. INTRODUÇÃO ..... **Erro! Indicador não definido.**
5.2. PROCESSO DE DESENVOLVIMENTO ..... **Erro! Indicador não definido.**
5.2.1. Ciclo do *Design Science Research* Aplicado ..... **Erro! Indicador não definido.**
5.2.2. Integração Metodológica e Validação ..... **Erro! Indicador não definido.**
5.3. CARACTERIZAÇÃO E COMPONENTES DO PTT ..... **Erro! Indicador não definido.**

5.3.1. Objetivos e Escopo ..... **Erro! Indicador não definido.**
5.3.2. Arquitetura Técnica ..... **Erro! Indicador não definido.**
5.3.3. Diferencial Inovador ..... **Erro! Indicador não definido.**
5.4. RESULTADOS DE IMPLEMENTAÇÃO E VALIDAÇÃO ... **Erro! Indicador não definido.**
5.4.1. Performance Quantitativa ..... **Erro! Indicador não definido.**
5.4.2. Análise Dimensional ..... **Erro! Indicador não definido.**
5.4.3. Validação da Sustentabilidade ..... **Erro! Indicador não definido.**
5.5. POTENCIAL DE REPLICABILIDADE E IMPACTOS ..... **Erro! Indicador não definido.**
5.5.1. Adaptabilidade Institucional ..... **Erro! Indicador não definido.**
5.5.2. Impactos Projetados ..... **Erro! Indicador não definido.**
5.5.3. Contribuição Científica ..... **Erro! Indicador não definido.**
5.6. DIRETRIZES DE IMPLEMENTAÇÃO ..... **Erro! Indicador não definido.**
5.6.1. Pré-requisitos Institucionais ..... **Erro! Indicador não definido.**
5.6.2. Fases de Implementação ..... **Erro! Indicador não definido.**
5.6.3. Governança e Monitoramento ..... **Erro! Indicador não definido.**
5.7. CONSIDERAÇÕES FINAIS ..... **Erro! Indicador não definido.**
REFERÊNCIAS ..... **Erro! Indicador não definido.**
APÊNDICE A - PRODUTO TÉCNICO-TECNOLÓGICO 114
Sistema de *Prompts* Estruturados para Geração Automatizada de Termos de Referência em Dispensas de Licitação: Guia de Implementação para o Poder Judiciário 114

1 INTRODUÇÃO INTEGRATIVA

Neste primeiro capítulo são apresentados os aspectos iniciais que direcionam e são base para esta pesquisa como um todo. Para tanto, apresenta-se a contextualização da temática, posteriormente os objetivos e, em seguida, a justificativa sobre a relevância do estudo. Ao final do capítulo é detalhada a estrutura da dissertação, pois ela está estruturada na forma de ensaios, um para cada objetivo específico.

1.1 CONTEXTUALIZAÇÃO

Os avanços recentes em Inteligência Artificial (IA) foram impulsionados por inovações em poder computacional, disponibilidade de dados e algoritmos de aprendizado de máquina, permitindo que sistemas aprendam a partir dos dados (Russell; Norvig, 2013). Ao realizar tarefas como classificação, previsão e identificação de fala e imagens, tem transformado o desempenho da sociedade (Bughin et al., 2018), contudo, à medida que a IA se torna cada vez mais presente, mais importante vão se tornando os aspectos de governança e de ética (Floridi et al., 2021), tanto no setor privado como no setor público.

A IA, e mais especificamente a Inteligência Artificial Generativa (IAG), representa um campo de inovação em rápida evolução, com significativo potencial para transformar a administração pública brasileira. Esta tecnologia, caracterizada pela capacidade de criar conteúdos originais a partir de padrões identificados em dados de treinamento, oferece oportunidades para otimizar procedimentos administrativos, reduzir a burocracia e promover maior eficiência na gestão dos recursos (Takale; Mahalle; Sule, 2024), inclusive públicos.

No contexto nacional, diversas iniciativas governamentais têm fomentado a transformação digital do setor público. O Programa de Transformação Digital do Governo Federal destaca-se como uma política estruturante, visando desburocratizar e modernizar a prestação de serviços através da incorporação de tecnologias emergentes (Brasil, 2020). Paralelamente, no âmbito do Poder Judiciário, o Conselho Nacional de Justiça (CNJ) implementou o Programa Justiça 4.0, desenvolvido em parceria com o Programa das Nações Unidas para o Desenvolvimento (PNUD), com o objetivo de incrementar a digitalização dos processos judiciais e administrativos (CNJ, 2024b; Ferraz, Caracas, Baggio, 2022).

Uma das tecnologias mais promissoras desse programa é a Plataforma Sinapses, primeira infraestrutura de inteligência artificial específica para o Judiciário brasileiro. Esta plataforma viabiliza o compartilhamento de projetos de IA entre tribunais, com cada órgão responsável por seus modelos e bases de dados, enquanto o CNJ mantém a infraestrutura

operacional (Maranhão, Junquilho, Tasso, 2023). Tal iniciativa ilustra o crescente interesse do sistema judiciário em soluções tecnológicas avançadas, embora o foco predominante tenha sido nas atividades jurisdicionais, com menor ênfase em funções administrativas como as licitações.

A escolha da modalidade de Dispensa de Licitação como escopo desta pesquisa fundamenta-se em sua relevância quantitativa e qualitativa nos processos de contratação dos Tribunais de Justiça, a exemplo do Tribunal de Justiça de Pernambuco (TJPE). Segundo dados do Portal da Transparência do TJPE, entre 2022 e 2024 esta modalidade constitui 17,64% das aquisições realizadas pelo TJPE. Some-se a isso o fato desta modalidade possuir características que a tornam especialmente adequada para a implementação experimental de IAG. A padronização relativa dos documentos exigidos nesta modalidade oferece um cenário controlado para a aplicação de tecnologias generativas, com potencial para maximizar a eficiência sem comprometer a segurança jurídica (Torres, 2021).

A opção específica pela Compra Direta, uma das hipóteses previstas no art. 75 da Lei nº 14.133/2021, delimita um recorte onde a automação pode proporcionar benefícios expressivos. Embora simplificados em comparação com outras modalidades licitatórias, esses procedimentos ainda exigem rigor técnico e conformidade legal, configurando um campo fértil para a aplicação de IAG na geração de documentos administrativos (Alencar; Bliacheriene; Araújo, 2024).

Diante desse cenário propício, especialmente na esfera do Poder Judiciário, compreender as ferramentas de IAG mais empregadas e seu impacto na transformação dos processos administrativos torna-se essencial. Nesse contexto, destaca-se o ChatGPT, desenvolvido pela OpenAI, cuja versatilidade, precisão textual e ampla adoção institucional justificam sua escolha como ferramenta central deste estudo.

A relevância dessa escolha é corroborada pelo Relatório de Pesquisa sobre o uso da IAG no Judiciário, publicado pelo Conselho Nacional de Justiça (CNJ, 2024e), o qual evidencia que o ChatGPT figura entre as plataformas mais utilizadas por tribunais brasileiros, tanto em projetos-piloto quanto em protótipos voltados à automatização e análise textual. Sua adoção por instituições como o TJMG, TJDFT, TJSP e o próprio TJPE reforça a legitimidade e a aplicabilidade do presente trabalho, alinhando-o não apenas às tendências tecnológicas atuais, mas também às práticas já consolidadas no sistema de justiça.

Contudo, ainda que promissora, a implementação de sistemas baseados em IAG no contexto das licitações públicas impõe desafios significativos que não podem ser negligenciados. Souza *et al*. (2024), em um mapeamento sistemático da literatura sobre o uso da IA no judiciário brasileiro, identificaram que a falta de transparência e explicabilidade

representa a preocupação mais recorrente, com 31 citações em 45 artigos analisados. Em seguida, destacam-se os vieses algorítmicos, mencionados em 23 estudos, apontando para riscos que comprometem diretamente a confiabilidade das soluções baseadas em IA.

Entre os principais riscos, destaca-se o fenômeno das chamadas "alucinações" – geração de informações aparentemente plausíveis, mas factualmente incorretas –, que representa um obstáculo crítico, especialmente em ambientes onde a precisão jurídica é imperativa (Tonmoy et al., 2024). No domínio das licitações públicas, esse problema pode resultar na invalidação de procedimentos ou em decisões juridicamente insustentáveis, comprometendo a legalidade e a eficiência administrativa. Estudos recentes demonstram que as alucinações constituem uma limitação inerente aos atuais Grandes Modelos de Linguagem (LLMs), exigindo, portanto, o desenvolvimento de estratégias específicas de mitigação (Xu et al., 2024).

Além disso, como já mencionado, os vieses algorítmicos constituem outra preocupação relevante. A dependência excessiva de sistemas automatizados na elaboração de documentos licitatórios pode, inadvertidamente, perpetuar preconceitos presentes nos dados de treinamento, resultando em discriminações indiretas ou na imposição de barreiras injustas a determinados grupos de fornecedores (Cambi; Amaral, 2023). Somam-se a isso os riscos decorrentes da chamada "caixa-preta algorítmica", ou seja, a opacidade dos mecanismos de funcionamento dos modelos, o que pode comprometer os princípios da transparência e da equidade – fundamentos basilares da administração pública (Fornasier; Silva; Schwede, 2023).

É neste cenário complexo que a engenharia de *prompts* emerge como um campo crucial para viabilizar a aplicação segura e eficaz da IAG. A formulação precisa de instruções para modelos generativos pode reduzir significativamente a ocorrência de alucinações e otimizar a qualidade dos artefatos produzidos. Pesquisas recentes evidenciam que técnicas avançadas de engenharia de *prompts* permitem estruturar melhor as respostas das IAGs, possibilitando um raciocínio mais elaborado e contextualmente alinhado (Schulhoff et al., 2024). No âmbito das licitações públicas, estas técnicas podem ser determinantes para a geração de documentos juridicamente precisos e administrativamente eficientes.

Ademais, a motivação para esta pesquisa transcende seu potencial teórico, ancorando-se também em sua aplicabilidade prática no TJPE. Como servidora atuante na área de licitações desta instituição, a pesquisadora tem observado diretamente os desafios operacionais relacionados à elaboração de documentos licitatórios, especialmente quanto à conformidade legal e à eficiência administrativa. A Secretaria de Administração, à qual está subordinado o setor de licitações do TJPE, tem demonstrado interesse crescente na aplicação de tecnologias de IA para otimizar processos, inclusive estabelecendo um comitê informal dedicado à

investigação dessas possibilidades. Este ambiente institucional receptivo à inovação tecnológica proporciona condições favoráveis para a condução de pesquisas aplicadas, permitindo não apenas o desenvolvimento teórico, mas também a validação prática das soluções propostas.

Assim, a presente pesquisa propõe-se a investigar como a engenharia de *prompts* pode ser estruturada para viabilizar o uso seguro, transparente e eficaz da IAG na elaboração de artefatos administrativos voltados à contratação direta por Dispensa de Licitação no TJPE. O estudo parte da premissa de que a implementação responsável dessa tecnologia pode contribuir significativamente para a modernização e eficiência da gestão pública, desde que observados os princípios éticos, legais e administrativos que regem a administração pública brasileira.

1.2 OBJETIVOS

Para implementar a pesquisa desejada, foram traçados os objetivos geral e específicos a seguir detalhados.

1.2.1 Objetivo Geral

Investigar como a engenharia de *prompt* pode ser estruturada para viabilizar o uso ético, transparente da Inteligência Artificial Generativa na elaboração de artefatos administrativos voltados à contratação direta por Dispensa de Licitação, a partir do estudo de caso do Tribunal de Justiça de Pernambuco.

1.2.2 Objetivos Específicos

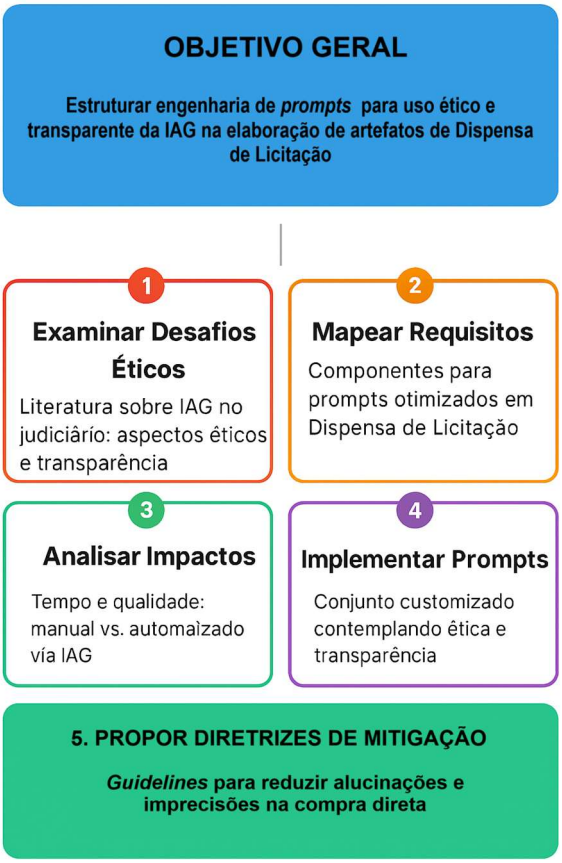
Para atingir o objetivo geral foram definidos os seguintes objetivos específicos:

1. Examinar na literatura científica os desafios e oportunidades da inteligência artificial generativa no judiciário brasileiro, com destaque para os aspectos éticos e de transparência;
2. Mapear os requisitos necessários para a estruturação de *prompts* otimizados para a geração de artefatos de licitação, na modalidade Dispensa de Licitação;
3. Analisar os impactos de tempo e qualidade da geração dos artefatos de licitação automatizada via inteligência artificial generativa;
4. Implementar um conjunto de *prompts* customizados e validados para geração de artefatos de licitação na modalidade de Dispensa de Licitação que contemplem os aspectos éticos e de transparência;

5. Propor diretrizes para mitigação de riscos relacionados a alucinações e imprecisões na geração automatizada de artefatos de licitação na modalidade Compra Direta.

A figura 1 sintetiza visualmente os principais componentes investigados nesta etapa, evidenciando a inter-relação entre os desafios éticos, a estrutura de *prompts* , as diretrizes propostas e os critérios de validação técnica, elementos que fundamentam a construção de soluções seguras e eficazes em engenharia de prompt.

Figura 1: Objetivos geral e específicos.



Fonte: Elaborado pela autora (2025).

1.3 JUSTIFICATIVA

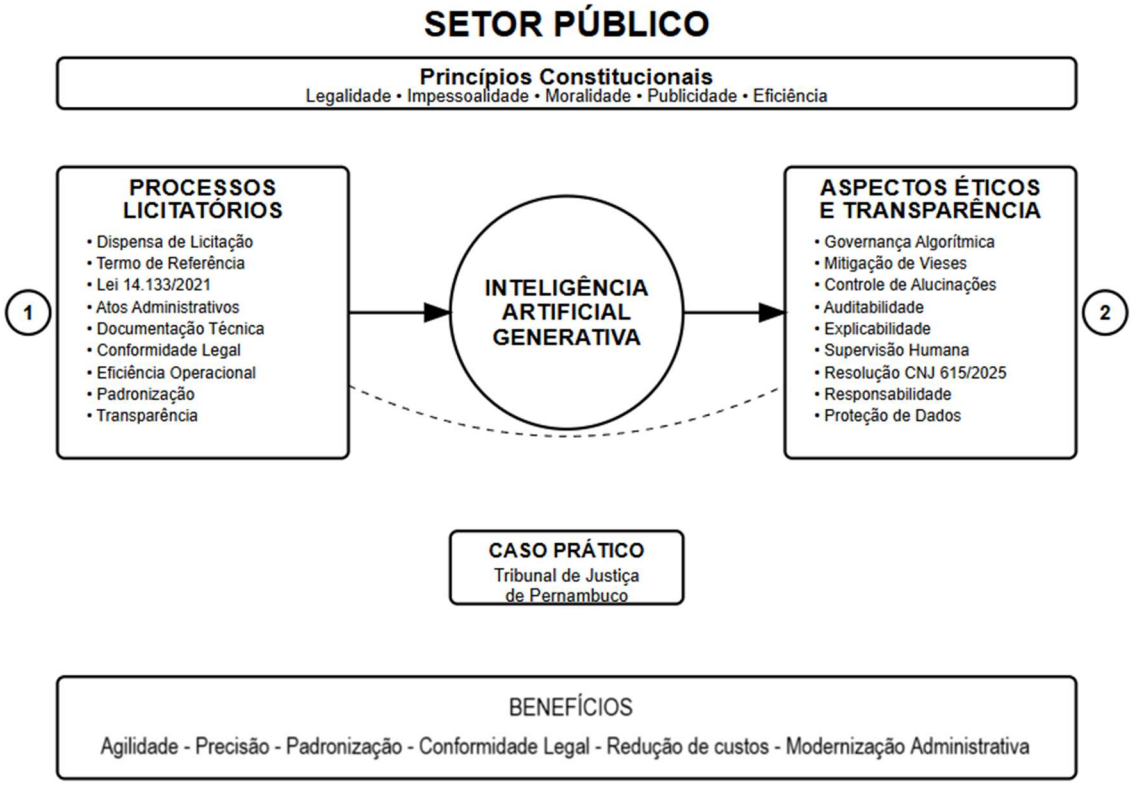
A crise do Poder Judiciário brasileiro e a dificuldade de lidar com o crescente volume de litígios têm sido objeto de discussão contínua no meio acadêmico e jurídico. O aumento exponencial de ações pendentes de julgamento, somado à demora na prestação jurisdicional, tem gerado insatisfação generalizada na sociedade (Tomasevicius Filho, 2015).

Essa situação crítica é amplamente reconhecida na literatura especializada, onde há consenso de que o modelo tradicional de ampliação das estruturas judiciais para lidar com a crescente demanda de litígios não é mais viável, especialmente em um contexto de crise

econômica (Brito, Fernandes, 2020). Além disso, o cenário pós-pandemia de COVID-19 reforçou a necessidade de adotar soluções inovadoras, já que o uso de tecnologias no âmbito jurisdicional se consolidou como um caminho sem volta

A crescente convergência entre modernização administrativa e inovação tecnológica no setor público brasileiro materializa-se de forma emblemática na aplicação da Inteligência Artificial Generativa aos processos licitatórios. Esta intersecção, que combina o imperativo da eficiência operacional com a rigorosa observância aos princípios constitucionais da administração pública, pode ser compreendida através de uma abordagem sistêmica que articula tecnologia, marco normativo e contexto institucional, conforme ilustrado na figura 2.

Figura 2: Inteligência Artificial Generativa nos processos licitatórios do setor público



Fonte: Elaborado pela autora (2025).

A trajetória de modernização tecnológica do Poder Judiciário brasileiro pode ser compreendida através de fases evolutivas distintas, cada uma caracterizada por marcos específicos e diferentes níveis de impacto transformacional. Esta evolução, que se intensificou dramaticamente a partir de 2020, evidencia uma progressão natural que culmina na atual fase de implementação da Inteligência Artificial, conforme sistematizado no Quadro 1.

A busca por alternativas tecnológicas, como a IA, surge como uma resposta para aliviar os gargalos do Judiciário, proporcionando maior agilidade e eficiência na tramitação dos processos. A aplicação de IAG também tem se destacado como uma das tecnologias mais promissoras para a automação de processos complexos e a otimização de tarefas administrativas no setor público (Desordi, Della Bona, 2020). No contexto do Judiciário brasileiro, que lida com uma grande quantidade de documentos e processos burocráticos, a aplicação dessa tecnologia pode representar um avanço significativo em termos de eficiência, redução de custos e precisão na elaboração de artefatos, como por exemplo, de licitação.

Quadro 1: *Timeline* da Transformação Digital no Judiciário Brasileiro

Período	Características Principais	Marcos Relevantes	Nível de Impacto
1988-2019 Era Pré-Digital	• Processos majoritariamente físicos • Baixa integração tecnológica • Crescimento do congestionamento • Iniciativas isoladas de informatização	• Lei 11.419/2006 - Processo Eletrônico • Resolução CNJ 185/2013 - PJe • Primeiros sistemas de gestão	**PREPARATÓRIO**
2020-2021 Catalisador Pandêmico	• Digitalização emergencial acelerada • Audiências 100% virtuais • Trabalho remoto generalizado • Adaptação forçada às tecnologias	• Resoluções CNJ emergenciais • Regulamentação de audiências virtuais • Consolidação do processo eletrônico	**ACELERADOR**
2021-2024 Programa Justiça 4.0	• Transformação digital estruturada • 4 eixos estratégicos • Gestão baseada em dados • Plataformas integradas	• PDPJ-Br e Plataforma Sinapses • Juízo 100% Digital • Balcão Virtual • Núcleos de Justiça 4.0	**ESTRUTURAL**

Período	Características Principais	Marcos Relevantes	Nível de Impacto
2022-Presente Era da IA	• Crescimento de 271% em projetos de IA • 49% dos magistrados usam IAG • 140 projetos mapeados (2023) • Emergência da IAG	• Resolução CNJ 332/2020 • Resolução CNJ 615/2025 • TCU lança ChatTCU (2023) • 96,4% preferem ChatGPT	**REVOLUCIONÁRIO**
2025-2030 IAG Responsável	• IAG em processos administrativos • Governança algorítmica madura • Transparência e auditabilidade • Capacitação especializada	• Implementação da Res. 615/2025 • Padronização entre tribunais • Compliance com LGPD • Mitigação de vieses	**TRANSFORMACIONAL**

Fonte: Baseado em CNJ (2024a, 2024e), Ferraz, Caracas e Baggio (2022)

Contudo, para garantir que o tratamento e armazenamento de dados confidenciais ocorram com o mais alto nível de segurança e confidencialidade, a contratação de uma nuvem privada apresenta-se como uma solução preferível. Esse ambiente privado protege as informações institucionais contra vulnerabilidades e exposições associadas ao uso de plataformas públicas, permitindo que dados sensíveis sejam processados exclusivamente sob os protocolos de segurança e gestão internos. Esta preocupação é especialmente relevante considerando os riscos de segurança associados ao *prompting*, que Schulhoff *et al*. (2024) classificam como "*prompt hacking*", incluindo técnicas como *prompt injection* e *jailbreaking*, que podem comprometer a integridade e confidencialidade dos sistemas baseados em IA. O Quadro 2 apresenta a terminologia específica relacionada a essas vulnerabilidades de segurança.

Conforme demonstrado por Bai *et al.* (2022), as IAGs treinadas adequadamente podem gerar texto coerente e contextualmente relevante, o que, no âmbito das licitações, pode se traduzir na rápida elaboração de minutas de documentos complexos. Esta capacidade não apenas economiza tempo, mas também permite que servidores públicos concentrem seus

esforços em atividades de maior valor agregado, como a análise estratégica e a tomada de decisões críticas.

Quadro 2: Terminologia de Segurança em *Prompting*

Termo	**Definição**	**Características**
Prompting	Processo de formulação de instruções específicas para modelos de IA generativa	Técnica fundamental para orientar o comportamento de LLMs sem modificar parâmetros do modelo
Prompt Hacking	Categoria de técnicas maliciosas que exploram vulnerabilidades em sistemas baseados em *prompts*	Visa manipular comportamento da IA, contornar salvaguardas ou extrair informações sensíveis
Prompt Injection	Inserção de entradas maliciosas em *prompts* para subverter comportamento pretendido	Análogo aos ataques de injeção SQL, mas aplicado a modelos de linguagem
Jailbreaking	Método para contornar restrições éticas e de segurança de modelos de IA	Manipula *prompts* para fazer o modelo ignorar diretrizes de segurança

Fonte: Baseado em Schulhoff *et al.* (2024).

Adicionalmente, a implementação de IAG nos processos licitatórios pode contribuir para a padronização e consistência dos documentos produzidos em diferentes órgãos e esferas da administração pública. Esta uniformização tem o potencial de reduzir erros, facilitar a interpretação por parte dos licitantes e órgãos de controle, e promover uma maior isonomia nos processos de contratação.

Contudo, como alertado por Contini (2024), é fundamental que essa padronização seja implementada de forma cuidadosa, garantindo que as especificidades de cada contratação sejam adequadamente consideradas e que não haja uma homogeneização excessiva que possa limitar a inovação ou a adaptabilidade às necessidades locais.

No entanto, Xu, Jain e Kankanhalli (2024) ressaltam que é crucial reconhecer as limitações inerentes destes sistemas, particularmente no que tange ao fenômeno das alucinações, e implementar mecanismos robustos de verificação e validação humana para garantir a integridade e a legalidade das avaliações realizadas. Schulhoff *et al*. (2024) corroboram esta preocupação e apresentam técnicas específicas para mitigar riscos de alucinação, como o uso de agentes baseados em LLMs equipados com ferramentas externas para verificação factual, e métodos de avaliação que aplicam *chain-of-thought* para melhorar a precisão das saídas geradas.

Chain-of-thought (cadeia de pensamento) representa uma técnica avançada de engenharia de *prompts* que se fundamenta na decomposição do raciocínio em etapas sequenciais e logicamente encadeadas, permitindo que modelos de linguagem processem

problemas complexos de maneira estruturada e verificável. Conforme proposto por Wei *et al*. (2022), esta abordagem transcende os métodos convencionais de *prompting* ao induzir o modelo a articular explicitamente cada estágio intermediário do processo cognitivo, estabelecendo uma trilha de raciocínio que conecta premissas e conclusões de forma transparente.

No contexto jurídico-administrativo das licitações públicas, a técnica *chain-of-thought* apresenta valor excepcional ao estruturar o raciocínio normativo de forma alinhada aos princípios de fundamentação e motivação dos atos administrativos, facilitando a produção de documentos que não apenas oferecem conclusões tecnicamente precisas, mas também explicitam o percurso lógico-jurídico que as sustenta. Liu *et al*. (2023) observam que esta técnica potencializa significativamente a capacidade dos LLMs em tarefas que exigem múltiplas operações lógicas interdependentes, enquanto Sahoo *et al*. (2024) evidenciam que a abordagem não apenas melhora a precisão dos resultados, mas também amplifica a interpretabilidade e auditabilidade das saídas geradas – características particularmente valiosas em procedimentos licitatórios, onde a justificação das decisões é tão importante quanto seu conteúdo substantivo.

Apesar desse conjunto extenso de possibilidades e da IAG ter sido amplamente usada no apoio à decisão judicial, não têm sido observadas muitas aplicações em atividades administrativas do judiciário. Lançado em 2023, o TCU foi o pioneiro no desenvolvimento de uma ferramenta chamada ChatTCU (TCU, 2024a), que é um assistente virtual de inteligência artificial generativa baseado no ChatGPT, desenvolvido pela OpenAI[1].

Esse *chatbot*[2] utiliza processamento de linguagem natural e foi criado para otimizar tarefas como a produção de textos, adaptações para linguagem simples, traduções e análises relacionadas às ações de controle externo do Tribunal (TCU, 2024). Entretanto, sua execução ocorre na nuvem da Microsoft, em um ambiente privado. As conversas são protegidas por contrato, assegurando a confidencialidade das informações fornecidas pelos usuários.

---

[1] OpenAI é uma empresa de pesquisa em inteligência artificial fundada em dezembro de 2015 por um grupo de empreendedores, pesquisadores e filantropos, incluindo Elon Musk e Sam Altman, cuja missão declarada é "garantir que a inteligência artificial geral (IAG) - sistemas de IA altamente autônomos que superam os humanos na maioria das tarefas economicamente valiosas - beneficie toda a humanidade" (OpenAI, 2024).

[2] *Chatbots* (termo derivado de "chat", que significa conversa, e "*bot*", abreviação de robô) são programas que visam imitar a interação humana durante uma conversa com usuários. Esses sistemas têm como finalidade responder perguntas de maneira que o usuário sinta que está dialogando com uma pessoa, em vez de um software. Após o usuário enviar uma pergunta via mensagem de texto, o *chatbot* busca informações em sua base de dados ou na internet e, em seguida, devolve uma resposta em uma linguagem clara e acessível. Dessa forma, o programa gera respostas automáticas a partir de palavras-chave identificadas na consulta. Além disso, *chatbots* utilizam técnicas de aprendizado de máquina para aprimorar a precisão das respostas, tornando as interações mais naturais e semelhantes às conversas humanas (Andrade, 2020, p. 31).

A escolha do ChatGPT como ferramenta base para esta investigação fundamenta-se em evidências empíricas robustas sobre seu uso predominante no contexto judiciário brasileiro. Conforme demonstrado pelo relatório do CNJ sobre o uso de IAG no Poder Judiciário, e evidenciado na figura 22, o ChatGPT é utilizado por 96,4% dos magistrados e 94,5% dos servidores que já tiveram experiência com tecnologias de IAG, consolidando-se como a ferramenta de referência no setor (CNJ, 2024e). Esta supremacia não se justifica apenas pela popularidade, mas pela capacidade superior da ferramenta em processar linguagem natural complexa e manter consistência em textos longos, características essenciais para a elaboração de documentos administrativos que demandam rigor técnico e conformidade legal (Sahoo et al., 2024). Ademais, a compatibilidade do ChatGPT com técnicas avançadas de engenharia de *prompts* , particularmente *Chain-of-Thought* e *Self-Refine*, oferece o ambiente tecnológico adequado para desenvolver soluções que atendam aos requisitos específicos das contratações diretas no âmbito do TJPE (Schulhoff et al., 2024).

Nesse contexto, a adoção de IAGs nas licitações públicas alinha-se com as tendências globais de digitalização e modernização da administração pública, potencialmente aumentando a atratividade e a competitividade dos certames. Todavia, como apontado por Morison e McInerney (2024), é imperativo que a implementação destas tecnologias seja acompanhada de medidas que garantam a transparência, a explicabilidade e a equidade dos processos automatizados, preservando assim os princípios fundamentais que regem as contratações públicas.

Assim, a justificativa para a realização deste estudo baseia-se em três fatores principais: a necessidade de modernização e automação dos processos judiciais, a possibilidade de aprimorar a gestão pública e os desafios éticos e jurídicos inerentes ao uso de IAG no setor público, como será detalhado a seguir.

1.3.1 Aprimoramento da Gestão Pública e Modernização do Judiciário

A aplicação da IAG no setor público pode gerar impactos positivos na gestão de recursos. A automação de tarefas burocráticas, como a criação de termos de referência e editais, libera profissionais para atividades de maior complexidade e valor estratégico. Além disso, a capacidade da IA de seguir normas e regulamentos específicos reduz a incidência de falhas e retrabalhos, otimizando o uso dos recursos públicos e promovendo maior transparência e conformidade legal nos processos administrativos.

Ademais, a adoção de tecnologias de IA na administração pública pode proporcionar economias significativas e melhorar a qualidade dos serviços oferecidos aos cidadãos (Barbosa,

2023). Portanto, a implementação de IAG no Judiciário alinha-se com as metas de eficiência operacional e responsabilidade fiscal da gestão pública.

Atualmente, o Judiciário brasileiro enfrenta o desafio de lidar com um grande volume de processos que, frequentemente, resulta em atrasos e ineficiências administrativas. Os processos licitatórios, em especial, exigem rigor técnico e jurídico, o que demanda tempo e recursos humanos.

A implementação da IAG para a criação de artefatos de licitação pode automatizar etapas críticas desses processos, oferecendo maior velocidade na elaboração de documentos, redução de erros humanos e maior padronização das informações. Conforme pesquisas recentes, a IA tem o potencial de transformar a maneira como o Judiciário lida com essas tarefas, promovendo uma gestão pública mais eficiente (Rossetti; Garcia, 2023).

1.3.2 Desafios Éticos e Jurídicos

Apesar dos benefícios, o uso de IAG no Judiciário levanta questões fundamentais sobre ética e responsabilidade legal, especialmente em relação à veracidade e à fidelidade das informações geradas pelos modelos de IA. O fenômeno das alucinações, em que a IA gera conteúdo factualmente incorreto ou sem base nas informações fornecidas, pode comprometer a integridade dos processos judiciais (Xu et al., 2024).

Schulhoff *et al*. (2024) identificam questões críticas de segurança e alinhamento relacionadas ao uso de técnicas de *prompting*, incluindo riscos de vazamento de dados de treinamento, exposição de *templates* de *prompt* considerados propriedade intelectual, e a geração de código com vulnerabilidades de segurança. No contexto de licitações públicas, onde documentos frequentemente contêm especificações técnicas detalhadas e cláusulas contratuais sensíveis, esses riscos são particularmente relevantes e exigem a implementação de salvaguardas adequadas, como defesas baseadas em *prompt*, *guardrails*[3] e sistemas de detecção de entradas maliciosas.

[3] *Guardrails* (barreiras de proteção) são mecanismos e estruturas de controle implementados em sistemas de Inteligência Artificial Generativa para impor restrições operacionais, prevenir comportamentos indesejados e garantir que os modelos funcionem dentro de parâmetros éticos e legais predefinidos. Conforme explicado por Schulhoff et al. (2024), estes dispositivos de segurança atuam como salvaguardas algorítmicas que filtram, monitoram e restringem as saídas dos modelos, assegurando conformidade normativa e mitigando riscos específicos como vazamento de dados sensíveis, geração de conteúdo inadequado ou manifestação de vieses. Essa prática se coaduna com as orientações de governança algorítmica propostas pelo CNJ (2020) e com as recomendações da literatura especializada, que destacam a necessidade de implementar barreiras técnicas e normativas para mitigar riscos associados a vieses, alucinações e violações de princípios éticos em soluções de IA (CONTINI, 2024; LIU et al., 2023; CNJ, 2020).

Além disso, questões como a propriedade intelectual e a responsabilidade jurídica sobre os documentos gerados pela IA precisam ser abordadas de maneira criteriosa (Rossetti; Garcia, 2023). Os desafios de alinhamento destacados por Schulhoff *et al*. (2024), como a sensibilidade excessiva a pequenas variações no *prompt* e a tendência dos modelos a exibir vieses presentes nos dados de treinamento, requerem atenção especial no desenvolvimento de sistemas para o setor público, onde a imparcialidade e a equidade são princípios fundamentais.

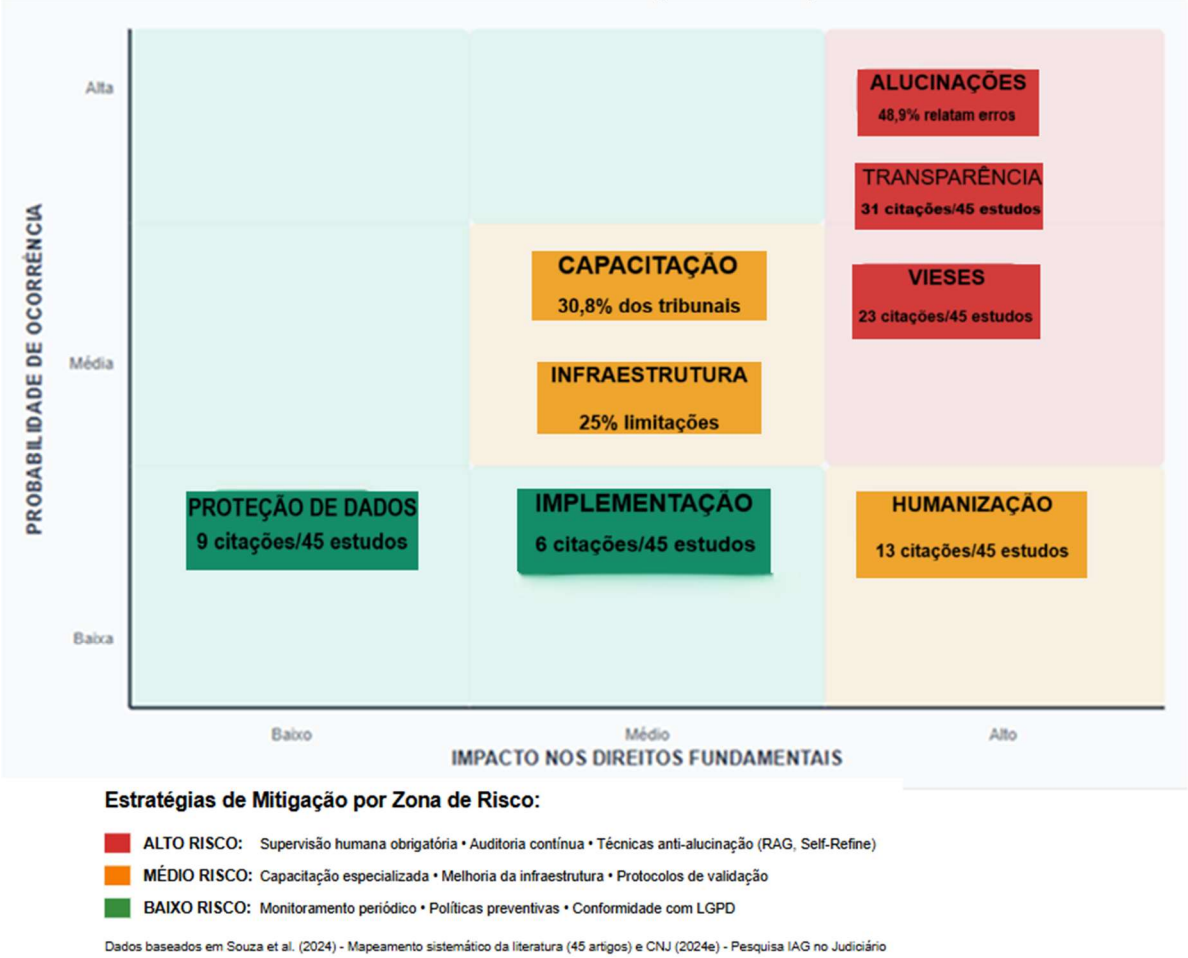
Diante disso, este estudo busca, além de explorar as possibilidades de otimização, propor soluções para mitigar os riscos associados ao uso da IAG, garantindo que sua implementação seja segura, eficaz e ética no contexto do Judiciário. O desenvolvimento de políticas públicas e diretrizes específicas para o uso responsável de IA é essencial para assegurar que essa tecnologia seja aplicada de maneira alinhada aos princípios de justiça, transparência e segurança jurídica (Costa et al*.,* 2023).

A magnitude desses desafios é amplificada pelo *déficit* democrático observado na governança tecnológica atual, que contrasta frontalmente com os princípios constitucionais de publicidade e participação que deveriam orientar a administração pública brasileira. O Painel CNJ 2024 corrobora essa preocupação ao revelar que menos de 15% dos projetos de IA no judiciário incluem mecanismos formais de consulta pública ou participação da sociedade civil em sua concepção e implementação.

Nesse contexto, a complexidade e multiplicidade dos riscos identificados na literatura especializada demandam uma abordagem analítica que permita compreender sua inter-relação e priorizar estratégias de mitigação baseadas em evidências empíricas, conforme sistematizado na figura 3.

A análise matricial evidencia que os desafios relacionados à transparência algorítmica, alucinações e vieses discriminatórios concentram-se na zona de alto risco, demandando intervenções prioritárias. Esta constatação fundamenta a ênfase desta pesquisa no desenvolvimento de técnicas específicas de engenharia de *prompts* que mitiguem sistematicamente essas vulnerabilidades críticas.

Figura 3: Desafios de implementação do IA no judiciário
Matriz de Riscos e Estratégias de Mitigação



Fonte: Baseado em Souza *et al.* (2024) e CNJ (2024e).

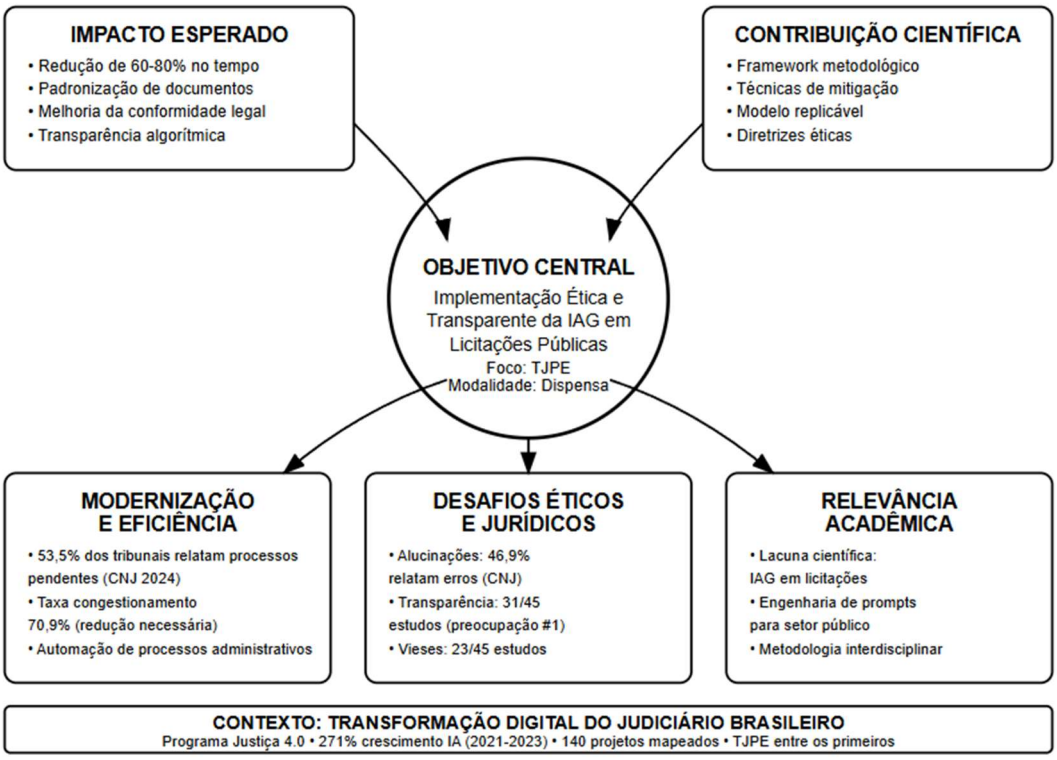
1.3.3 Relevância Acadêmica e Social

A pesquisa sobre o uso de IAG no Judiciário também possui grande relevância acadêmica e social. Do ponto de vista acadêmico, o estudo contribuirá para o desenvolvimento de novas abordagens teóricas e metodológicas sobre a aplicação de tecnologias emergentes na administração pública, promovendo uma maior compreensão dos impactos da IAG no setor jurídico. Socialmente, os benefícios da automação podem resultar em um Judiciário mais ágil e acessível, o que melhora a prestação de serviços à sociedade como um todo.

A convergência entre os imperativos de modernização administrativa, os desafios éticos emergentes e a relevância científica da temática constitui o tripé fundamental que sustenta a

justificativa desta pesquisa. A articulação desses três pilares, no contexto específico da transformação digital do Judiciário brasileiro, evidencia não apenas a oportunidade, mas a necessidade premente de investigar a aplicação responsável da IAG em processos licitatórios, conforme sistematizado na figura 4.

Figura 4: Justificativa do projeto



Fonte: Elaborado pela autora (2025)

Como evidenciado na síntese visual, esta pesquisa posiciona-se na intersecção estratégica entre demanda operacional comprovada, urgência ética documentada e lacuna científica identificada, constituindo contribuição simultaneamente teórica e prática para o avanço responsável da inteligência artificial no setor público brasileiro.

1.4 ESTRUTURA DA DISSERTAÇÃO

A presente dissertação está organizada em quatro capítulos complementares, que abordam desde os fundamentos teóricos até os aspectos metodológicos e práticos do estudo, conforme detalhado a seguir. Vale salientar que a presente dissertação optou pelo formato de ensaios, ao invés de um texto corrido, como é o mais tradicional desse gênero textual.

Este primeiro capítulo, **Introdução Integrativa**, estabelece as bases da investigação por meio da contextualização da temática, delimitação dos objetivos e justificativa de sua relevância, proporcionando ao leitor uma visão abrangente do problema de pesquisa e sua significância.

O segundo capítulo, correspondente ao primeiro ensaio, **Inteligência Artificial no Judiciário Brasileiro: Desafios e Oportunidades acerca dos Aspectos Éticos e de Transparência**, examina o panorama atual da IA no contexto judiciário, analisando criticamente seus desdobramentos éticos e os requisitos de transparência algorítmica. Este segmento corresponde ao primeiro objetivo específico, investigando as bases conceituais necessárias para a implementação responsável da tecnologia.

No terceiro capítulo, correspondente ao segundo ensaio, **Inteligência Artificial Generativa nas Licitações Públicas: Requisitos para Estruturação dos *Prompts***, são mapeados os elementos essenciais para a construção de instruções otimizadas para modelos de IAG no contexto das contratações públicas, fornecendo subsídios técnicos para a fase experimental do estudo, para realizar o segundo objetivo específico.

O quarto capítulo, correspondente ao terceiro ensaio, **Inteligência Artificial Generativa nas Licitações Públicas: Um Estudo de Caso no Tribunal de Justiça de Pernambuco**, apresenta a metodologia e o desenho da investigação empírica, detalhando os procedimentos para implementação e validação dos *prompts* customizados para geração de artefatos de licitação, alinhando-se ao terceiro objetivo específico.

No formato de um Apêndice independente, foi realizado o **Produto Técnico-Tecnológico** resultante da pesquisa, explicitando sua aplicabilidade prática e potenciais contribuições para o aprimoramento dos processos licitatórios no TJPE, com ênfase nos aspectos de mitigação de riscos relacionados a alucinações e imprecisões algorítmicas, conforme preconizado no quarto objetivo específico.

Esta estrutura foi concebida para proporcionar um encadeamento lógico entre os componentes da pesquisa, partindo dos fundamentos teóricos e normativos para culminar em uma aplicação prática e inovadora, capaz de contribuir significativamente para a modernização e eficiência dos processos administrativos no âmbito do Poder Judiciário.

2. INTELIGÊNCIA ARTIFICIAL NO JUDICIÁRIO BRASILEIRO: DESAFIOS E OPORTUNIDADES ACERCA DOS ASPECTOS ÉTICOS E DE TRANSPARÊNCIA

2.1. INTRODUÇÃO

A crescente adoção da Inteligência Artificial (IA) e da Inteligência Artificial Generativa (IAG) no âmbito da administração pública e, especialmente, no Judiciário brasileiro reflete uma tendência global para a automação de processos e otimização de serviços. Entretanto, esta transformação não se limita apenas aos aspectos operacionais, mas suscita questões éticas e de transparência que ainda carecem de tratamento adequado. Conforme apontado por Souza *et al.* (2024), a falta de transparência e explicabilidade dos algoritmos - frequentemente referida como fenômeno da "caixa-preta" - representa a preocupação mais recorrente na literatura científica sobre IA no judiciário brasileiro, com 31 citações em 45 artigos analisados.

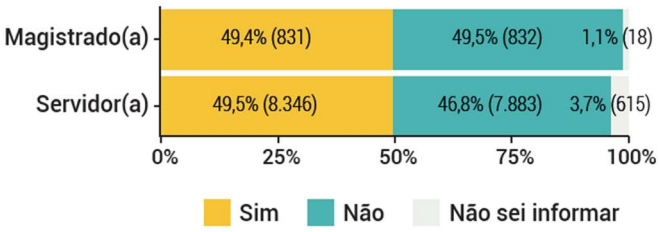
Os dados mais recentes do Conselho Nacional de Justiça (CNJ) consolidam essa preocupação e revelam uma expansão acelerada da IA no Judiciário. De acordo com a Pesquisa de Uso de IA no Poder Judiciário 2023, divulgada em maio de 2024 pelo presidente do CNJ, ministro Luís Roberto Barroso, foram identificados 140 projetos de IA desenvolvidos ou em desenvolvimento nos tribunais brasileiros - um crescimento de 26% em relação aos 111 projetos mapeados em 2022 (CNJ, 2024). Este crescimento expressivo, de 271% entre 2021 e 2023, evidencia tanto o potencial transformador da tecnologia quanto a urgência de estabelecer marcos éticos e de transparência adequados.

Como aponta o CNJ, o Poder Judiciário brasileiro enfrenta desafios crônicos de excesso de processos e morosidade, como evidenciado pelos 83,8 milhões de processos pendentes registrados pelo CNJ em 2023, dos quais 77% estão na Justiça Estadual (CNJ, 2024a). Este cenário de sobrecarga pressiona por soluções tecnológicas que, embora promissoras para aumentar a eficiência, não podem comprometer princípios fundamentais como imparcialidade, equidade e devido processo legal.

Contudo, a implementação da IAG no Judiciário brasileiro não está isenta de obstáculos. O relatório "O uso da Inteligência Artificial Generativa no Poder Judiciário Brasileiro", publicado pelo CNJ em setembro de 2024, revela aspectos preocupantes sobre o uso descentralizado e muitas vezes não transparente dessas tecnologias. A pesquisa, que envolveu 1.681 magistrados e 16.844 servidores, demonstrou que aproximadamente 50% dos

respondentes já utilizaram ferramentas de IAG (figura 5), mas a maioria (cerca de 85%) não informa sobre esse uso em suas atividades profissionais (CNJ, 2024e).

Figura 5: Percentual de uso de IAG



Fonte: CNJ (2024e, p. 52)

Essa falta de transparência representa um risco significativo para a governança ética da IA no Judiciário. Como destaca o relatório do CNJ, "a falta de transparência quanto ao uso entre os pares de uma organização, no caso, entre magistrados(as) e servidores(as), pode afetar negativamente os processos de revisão interna sobre o conteúdo do trabalho produzido na prestação do serviço jurisdicional" (CNJ, 2024e, p. 49).

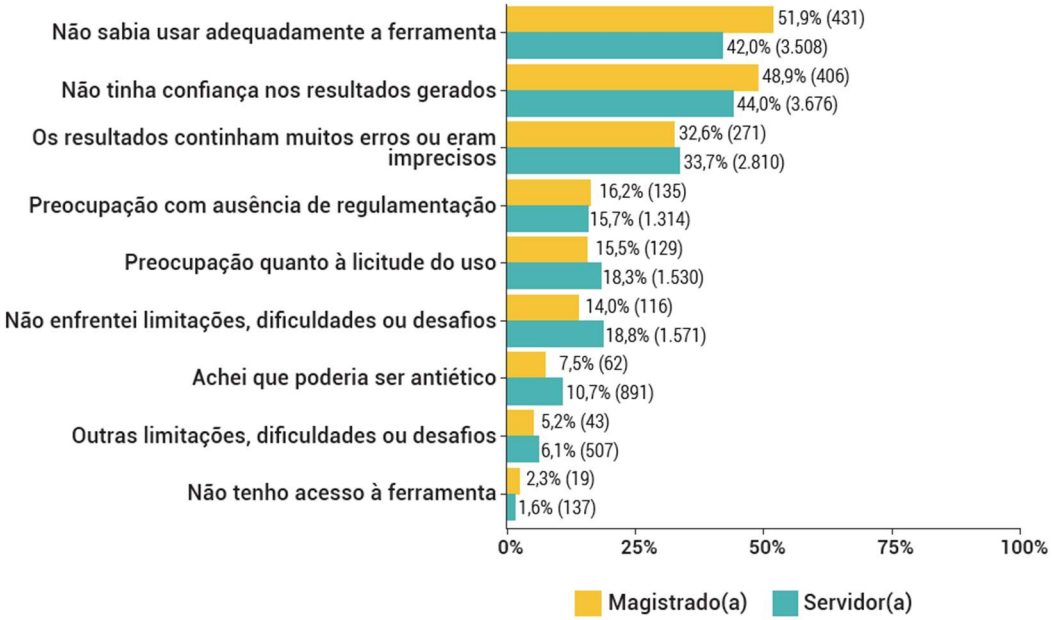
Tais obstáculos também foram abordados por Limberger, Giannakos e Szinvelski (2020) ao ressaltarem que, embora as ferramentas de IA possam otimizar o desempenho judiciário, questões técnicas e éticas, como a confiabilidade das decisões automatizadas e a necessidade de constante supervisão humana, são aspectos críticos que precisam ser cuidadosamente considerados. Esta observação ganha particular relevância quando confrontada com a realidade de implementação acelerada dessas tecnologias no judiciário brasileiro, muitas vezes sem o correspondente debate ético e normativo que deveria precedê-la.

A questão não se limita apenas à capacidade técnica da IAG, mas abrange também dimensões de responsabilidade jurídica, proteção de direitos fundamentais e governança algorítmica. Os riscos de vieses algorítmicos, citados em 23 dos 45 artigos analisados por Souza *et al.* (2024), evidenciam como sistemas aparentemente neutros podem perpetuar e até amplificar discriminações existentes, comprometendo a legitimidade e a justiça das decisões.

Uma descoberta particularmente relevante do relatório CNJ 2024 refere-se ao fenômeno das "alucinações" em sistemas de IAG. A pesquisa revelou que 48,9% dos magistrados indicaram "falta de confiança nos resultados gerados" e 32,6% relataram que "os resultados continham muitos erros", evidenciando a percepção dos usuários sobre limitações críticas da tecnologia (CNJ, 2024e), como sintetiza a figura 6. Esse dado é especialmente preocupante

quando confrontado com o uso significativo de IAG para pesquisas jurídicas: cerca de 15% dos magistrados e percentual significativo de servidores utilizam essas ferramentas para busca de jurisprudência, precedentes e leis.

Figura 6: Principais limitações, dificuldades ou desafios ao se utilizar ferramentas de IAG



Fonte: (CNJ, 2024e, p. 64)

A compreensão destas questões fornecerá subsídios essenciais para o desenvolvimento do quarto objetivo específico da dissertação, que busca "propor diretrizes para mitigação de riscos relacionados a alucinações e imprecisões na geração automatizada de artefatos de licitação na modalidade Compra Direta".

Esta investigação sobre os aspectos éticos e de transparência corresponde especificamente ao primeiro objetivo específico da dissertação, que visa "examinar na literatura científica os desafios e oportunidades da inteligência artificial generativa no judiciário brasileiro, com destaque para os aspectos éticos e de transparência". A análise aprofundada destes aspectos estabelece as bases conceituais e normativas necessárias para, posteriormente, desenvolver os *prompts* customizados para geração de artefatos de licitação que contemplem adequadamente essas dimensões éticas.

Nesse contexto, este capítulo se propõe, portanto, a responder a seguinte pergunta de pesquisa: como a aplicação da IAG no Judiciário brasileiro tem abordado questões éticas e de transparência, identificando lacunas significativas no atual modelo de implementação e propondo diretrizes para uma incorporação mais responsável dessas tecnologias, que preserve os valores fundamentais do Estado Democrático de Direito. Os *insights* obtidos nesta análise

serão fundamentais para orientar o desenvolvimento dos "*prompts* customizados e validados para geração de artefatos de licitação na modalidade de Dispensa de Licitação que contemplem os aspectos éticos e de transparência", conforme estabelecido no terceiro objetivo específico desta dissertação.

Para tanto, este ensaio teórico está estruturado em cinco seções complementares: inicialmente, apresentamos na seção 2.2 uma contextualização dos conceitos fundamentais da Inteligência Artificial, destacando suas diferentes classificações e implicações éticas; na sequência, a seção 2.3 explora as tecnologias de Processamento de Linguagem Natural e Grandes Modelos de Linguagem (LLMs), que são a base tecnológica da IAG, com ênfase em suas potencialidades e limitações; a seção 2.4 analisa especificamente como a IA tem sido implementada no judiciário brasileiro, identificando tendências e preocupações emergentes; a seção 2.5 examina o processo de transformação digital do Poder Judiciário, incluindo iniciativas como o Programa Justiça 4.0 e a Plataforma Sinapses, sob a perspectiva da governança e transparência; por fim, a seção 2.6 sintetiza os principais desafios éticos e apresenta perspectivas para uma implementação mais equilibrada e responsável da IA no contexto judiciário nacional.

2.2. INTELIGÊNCIA ARTIFICIAL

A compreensão do conceito de IA é fundamental para a análise crítica de seus impactos no sistema judiciário. Mais que uma ferramenta tecnológica, a IA representa um paradigma que reorganiza as relações de poder e tomada de decisão, com implicações profundas para a administração da justiça.

A IA é uma área de pesquisa relativamente nova, cujo desenvolvimento remonta ao período pós-Segunda Guerra Mundial, com o teste de Turing, proposto por Alan Turing (1950), que buscava definir operacionalmente a inteligência. O teste consistia em avaliar se um computador poderia enganar um interrogador humano, fazendo-o acreditar que estava interagindo com outra pessoa, por meio de perguntas e respostas escritas (Russell; Norvig, 2013). Para que uma máquina seja capaz de passar nesse teste, é necessário que ela possua algumas capacidades fundamentais, como processamento de linguagem natural, representação de conhecimento, raciocínio automatizado e aprendizado de máquina, habilidades essas que compõem os principais domínios da IA, complementam os recém citados autores.

O termo "inteligência artificial" foi cunhado em 1956, durante a conferência de Dartmouth, nos Estados Unidos, pelo cientista John McCarthy. Este a definiu como a atribuição de capacidades cognitivas a máquinas, utilizando códigos de programação e circuitos elétricos

para possibilitar interações autônomas (McCarthy, 2007). A partir desse marco, diferentes abordagens teóricas emergiram, classificando a IA de acordo com suas características e aplicações. De maneira geral, as definições de IA podem ser agrupadas em quatro perspectivas principais (Russell, Norvig, 2013), conforme sistematizado no Quadro 3.

Quadro 3: Abordagens Conceituais da Inteligência Artificial

Abordagem	**Descrição**	**Ênfase**	**Exemplo Típico**
Pensar como humanos	Busca simular processos cognitivos humanos (memória, aprendizagem, percepção)	Psicologia cognitiva e neurociência	Modelos de redes neurais artificiais e arquiteturas cognitivas
Pensar racionalmente	Baseia-se na lógica formal e nas "leis do pensamento" aristotélicas	Filosofia e lógica matemática	Sistemas especialistas baseados em regras lógicas
Agir como humanos	Foca em reproduzir externamente o comportamento humano observável	Teste de Turing e comportamento imitativo	*Chatbots* conversacionais e assistentes virtuais
Agir racionalmente	Busca a ação ótima para maximizar objetivos em determinado ambiente	Teoria da decisão e otimização	Agentes autônomos e sistemas de recomendação

Fonte: Baseado em Russell e Norvig (2013).

A abordagem cognitiva da IA busca modelar os processos internos do pensamento humano, analisando como os indivíduos resolvem problemas e tomam decisões. Essa perspectiva se alinha à ciência cognitiva, que integra modelos computacionais da IA com experimentos psicológicos para construir teorias verificáveis sobre o funcionamento da mente humana (Russell, Norvig, 2013).

Já a tradição logicista, fundamentada na filosofia aristotélica, busca desenvolver sistemas capazes de realizar inferências corretas a partir de premissas formais. Entretanto, essa abordagem enfrenta desafios significativos, como a dificuldade de representar conhecimento incerto em notação lógica e a complexidade computacional envolvida na resolução de problemas com um grande volume de fatos, como apontam os recém citados autores.

Por outro lado, a abordagem do agente racional, amplamente adotada na pesquisa moderna em IA, define um sistema inteligente como aquele que age de forma a maximizar os resultados esperados, mesmo em cenários de incerteza. Diferentemente da perspectiva logicista, essa abordagem não se limita à inferência lógica, incorporando também mecanismos como aprendizado de máquina e otimização probabilística. Assim, um agente racional deve ser capaz

de perceber o ambiente, processar informações, adaptar-se a mudanças e tomar decisões baseadas em objetivos específicos (Pereira, Botelho, 2018).

Em termos de classificação, é muito frequente que a IA seja subdividida entre IA fraca e IA forte. A IA mais utilizada é a IA fraca, ou estreita, caracterizada pela especialização em tarefas delimitadas, o que inclui sistemas como assistentes virtuais, algoritmos de recomendação e reconhecimento de imagens (Lage, 2021; Shabbir, Anwer, 2015). Em contrapartida, a IA forte, ou de propósito geral, aspira a uma inteligência comparável à humana, capaz de atuar em múltiplos domínios e realizar qualquer tarefa intelectual, mas ainda permanece como um objetivo teórico, sem implementação prática efetiva (Surden, 2019; Stone et al*.,* 2016).

Quanto ao seu processo evolutivo, a IA pode ser segmentada em três estágios, conforme Kaplan e Haenlein (2019):

- ANI (*Artificial Narrow Intelligence*): IA estreita, utilizada em sistemas como Facebook (reconhecimento facial), Siri (assistência por voz) e Tesla (veículos autônomos);
- AGI (*Artificial General Intelligence*): IA forte, ainda em estágio teórico;
- ASI (*Artificial Superintelligence*): conceito especulativo, que ultrapassaria a inteligência humana.

Dessa forma, enquanto a IA estreita, que também pode ser chamada de IA fraca, já é amplamente empregada, a busca por uma IA de propósito geral ainda enfrenta desafios significativos, tanto do ponto de vista tecnológico quanto ético.

2.3. PROCESSAMENTO DE LINGUAGEM NATURAL E GRANDE MODELO DE LINGUAGEM

O Processamento de Linguagem Natural (PLN), área fundamental para o desenvolvimento da IAG, consiste no desenvolvimento de modelos computacionais para a realização de tarefas que dependem de informações expressas em alguma língua natural, como tradução e interpretação de textos, busca de informações em documentos e interfaces homem-máquina (Liu et al., 2023, Pereira, s.d). Esta área interdisciplinar conecta a ciência da computação, a linguística e a inteligência artificial, buscando capacitar máquinas a compreender, interpretar e gerar linguagem humana de forma contextualmente adequada.

Conforme Pereira (s.d.), a pesquisa em PLN está voltada essencialmente a três aspectos da comunicação em língua natural: som (fonologia), estrutura (morfologia e sintaxe) e significado (semântica e pragmática). Na visão deste autor, a fonologia relaciona-se ao reconhecimento dos sons que compõem as palavras; a morfologia reconhece as palavras em

termos das unidades primitivas que as compõem; a sintaxe define a estrutura de uma frase com base na forma como as palavras se relacionam; a semântica associa significado a uma estrutura sintática, enquanto a pragmática verifica se o significado associado é realmente o mais apropriado no contexto considerado.

A evolução do PLN tem sido marcada pelo desenvolvimento de gramáticas formais, análise sintática automática e, mais recentemente, pela aplicação de técnicas de aprendizado de máquina, especialmente as redes neurais profundas (Sahoo et al*.,* 2024). Esta evolução culminou nos modelos de linguagem de grande escala (LLMs), que representam o estado da arte em processamento de linguagem natural, demonstrando capacidades sem precedentes em tarefas como geração de texto, tradução, resposta a perguntas e raciocínio, complementam os recém citados autores.

Os LLMs, construídos sobre a arquitetura *Transformer* (Vaswani et al*.,* 2017), revolucionaram o campo do PLN ao substituir as arquiteturas tradicionais baseadas em redes neurais recorrentes. A principal inovação desta arquitetura está em seu mecanismo de *self-attention*, que permite ao modelo avaliar a importância de diferentes partes da entrada em paralelo, capturando eficientemente as dependências de longo alcance em um texto (Shen *et al*., 2023). Este avanço permitiu um salto qualitativo significativo na compreensão contextual e na geração de texto por sistemas computacionais.

A capacidade dos LLMs de capturar nuances linguísticas, entender contextos e gerar respostas coerentes e relevantes os torna ferramentas valiosas para aplicações administrativas, incluindo a elaboração e análise de artefatos de licitação. No entanto, como destacado por Xu *et al*. (2024), estes modelos enfrentam desafios significativos, como o fenômeno das alucinações, que podem comprometer a confiabilidade e precisão de suas saídas, especialmente em contextos que exigem rigor técnico e conformidade legal.

A técnica de engenharia de *prompts* , discutida por Sahoo *et al*. (2024), emerge como uma abordagem fundamental para estender as capacidades dos LLMs. Diferente do paradigma tradicional de aprendizado supervisionado, a engenharia de *prompts* permite que estes sistemas abordem tarefas específicas usando instruções cuidadosamente projetadas, sem a necessidade de modificar os parâmetros do modelo. Esta técnica tem se mostrado particularmente eficaz em cenários onde a disponibilidade de dados rotulados é limitada, oferecendo uma alternativa flexível e adaptável para a utilização de LLMs em diversos contextos.

No cenário jurídico-administrativo, a aplicação de PLN e, especificamente, das técnicas de engenharia de *prompts* com LLMs, apresenta oportunidades significativas para a otimização de processos, incluindo a elaboração e análise de documentos licitatórios. Porém, como

observado por Contini (2024), esta aplicação deve ser cuidadosamente calibrada para garantir a precisão, conformidade legal e transparência necessárias em procedimentos administrativos públicos.

A aplicação de LLMs no contexto judiciário brasileiro enfrenta ainda o desafio adicional da especificidade linguística e jurídica local. A maioria destes modelos é treinada predominantemente em dados em língua inglesa e com base em sistemas jurídicos distintos do brasileiro, o que levanta questões sobre sua adequação para analisar e interpretar documentos legais em português e dentro do arcabouço normativo nacional. Este problema, embora técnico em sua natureza, tem profundas implicações éticas relacionadas à precisão e equidade das análises realizadas por estes sistemas.

A evolução contínua das técnicas de PLN, especialmente no contexto dos LLMs, promete transformar significativamente a forma como se interage com sistemas computacionais e como estes sistemas podem auxiliar em tarefas complexas que requerem compreensão e geração de linguagem natural. No entanto, a realização plena deste potencial depende da superação de desafios persistentes, como as alucinações, e do desenvolvimento de abordagens que garantam a confiabilidade, imparcialidade e explicabilidade destes sistemas (Contini, 2024; Sahoo et al*.,* 2024; Tonmoy et al*.,* 2024; Xu et al*.,* 2024).

2.4. A INTELIGÊNCIA ARTIFICIAL E O JUDICIÁRIO BRASILEIRO

A crise do Poder Judiciário brasileiro, caracterizada pelo excesso de ações pendentes e pela demora na prestação jurisdicional (Brito; Fernandes, 2020), é um tema recorrente na academia e no meio jurídico. Em 2023, o Conselho Nacional de Justiça (CNJ) registrou 83,8 milhões de processos pendentes, 77% deles na Justiça Estadual. Desse total, 63,6 milhões estão em análise, excluídos 18,5 milhões de processos suspensos, dos quais 2,5 milhões aguardam precedentes obrigatórios, como repercussão geral no STF e recursos repetitivos no STJ.

O ingresso de 35,3 milhões de novos casos em 2023 representou um aumento de 9,4% em relação a 2022, consolidando-se como o maior volume da série histórica. Apesar de o Judiciário ter julgado 33,2 milhões de processos, um aumento de 11,3% em relação ao ano anterior, o acervo processual cresceu em 896 mil casos. No mesmo período, 1,7 milhões de processos foram reativados, principalmente por anulação de sentenças ou conflitos de competência, e o Índice de Produtividade da Magistratura (IPM) subiu 6,8%, superando 2.000 processos baixados por magistrado.

A implementação massiva de soluções de IA tem sido uma das estratégias adotadas para enfrentar essa crise. Os dados mais recentes do CNJ revelam uma expansão acelerada: de 41 projetos em 2021 para 140 projetos em 2023, representando um crescimento de 271% em apenas dois anos (CNJ, 2024). Desses 140 projetos mapeados, 63 já estão em uso ou aptos a serem utilizados, enquanto 46 estão em fase final de desenvolvimento.

O Judiciário estadual concentra o maior número de projetos (68), seguido pelas justiças Eleitoral (23), do Trabalho (20), Federal (14) e Superior (13). O Tribunal de Justiça do Rio Grande do Sul lidera em número de projetos com 12 iniciativas, seguido pelo Tribunal de Justiça do Ceará com 7 e São Paulo com 6 projetos (CNJ, 2024).

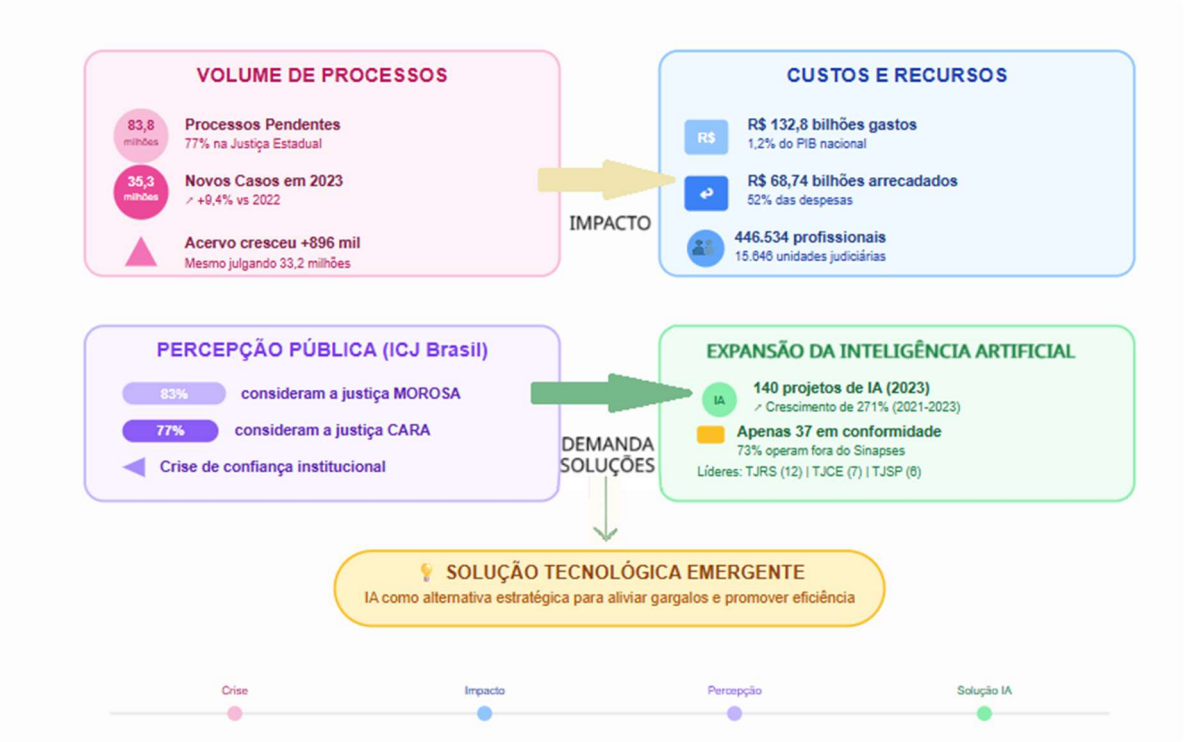
Entretanto, uma descoberta preocupante emerge da análise desses dados: dos 140 modelos mapeados, apenas 37 estão em conformidade e hospedados no Sinapses, a plataforma nacional do CNJ para IA no Judiciário (CNJ, 2024). Isso significa que 73% dos projetos de IA em desenvolvimento nos tribunais brasileiros operam fora dos marcos de governança estabelecidos pelo CNJ, evidenciando um grave problema de governança e supervisão institucional.

A magnitude desta crise estrutural e sua repercussão multidimensional podem ser visualizadas de forma integrada na figura 7, que articula o volume processual crítico, os custos operacionais, a percepção social negativa e a emergência de soluções tecnológicas baseadas em inteligência artificial como resposta estratégica aos gargalos identificados.

Com a implantação do Projeto Justiça 4.0 em 2021, que visa incrementar o processo de transformação digital do Poder Judiciário, houve melhorias significativas na eficiência. A taxa de congestionamento caiu para 70,5%, indicando que, de cada 100 processos, 30 foram baixados no ano de ingresso. O Índice de Atendimento à Demanda (IAD) chegou a 99,2%, quase equilibrando a entrada e saída de processos. No entanto, a estrutura judiciária, composta por 446.534 profissionais e 15.646 unidades, ainda não é suficiente para conter o crescimento da litigiosidade.

Das figuras 8 e 9 pode-se inferir que a implantação do Projeto Justiça 4 pode ter contribuído na melhoria de tais índices. A taxa de congestionamento caiu para 70,5%, o segundo melhor resultado em 15 anos. Isso significa que, a cada 100 processos tramitados, 30 foram baixados no mesmo ano de seu ingresso. Contudo, mesmo com esse esforço, o estoque de processos aumentou em 896 mil (CNJ, 2024)

Figura 7: Panorama da Crise do Poder Judiciário Brasileiro em 2023: volume processual, custos operacionais, percepção pública e emergência de soluções tecnológicas baseadas em Inteligência Artificial



Fonte: Baseado em CNJ (2024a) e Ramos *et al*. (2021).

Figura 8: Índice de produtividade por magistrado



Fonte: CNJ (2024, p. 20).

Figura 9: Índice de produtividade por servidor.

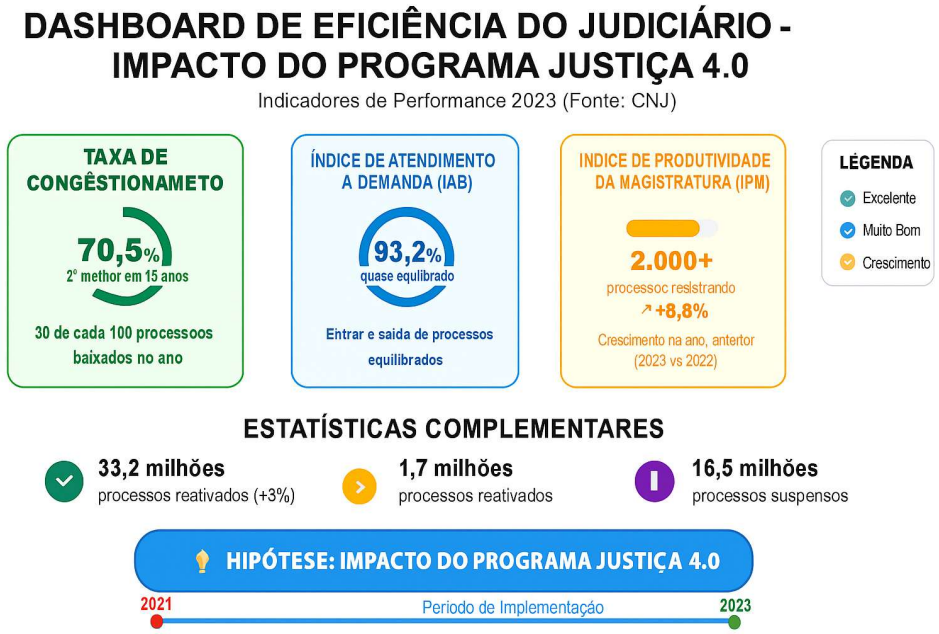


Fonte: CNJ (2024, p. 20).

O Judiciário gastou R$ 132,8 bilhões em 2023, o que corresponde a 1,2% do PIB nacional, arrecadando R$ 68,74 bilhões, equivalentes a 52% das despesas. Além disso, a percepção sobre a justiça é que ela é morosa (83%) e cara (77%), segundo os entrevistados do relatório ICJ Brasil (Ramos et al., 2021). Esse cenário reflete a crise de eficiência enfrentada pelo sistema, que persiste, apesar das reformas e da criação de microssistemas para julgamento de casos repetitivos (Brito; Fernandes, 2022).

Entretanto, os indicadores de performance consolidados na figura 10 evidenciam melhorias significativas nos principais índices de eficiência do Judiciário brasileiro, corroborando a hipótese de que a implementação do Programa Justiça 4.0 tem contribuído para a otimização operacional do sistema, conforme demonstrado pela redução da taxa de congestionamento para 70,5% e pelo incremento de 6,8% no Índice de Produtividade da Magistratura.

Figura 10: *Dashboard* de Indicadores de Eficiência do Poder Judiciário Brasileiro (2023): impacto do Programa Justiça 4.0 na Taxa de Congestionamento, Índice de Atendimento à Demanda e Índice de Produtividade da Magistratura.



Fonte: Baseado em CNJ (2024a).

Assim, a crise econômica e os desafios impostos pelo período pós-pandemia destacam a necessidade de soluções inovadoras. O uso de tecnologia, especialmente IAG, surge como uma alternativa relevante para otimizar o trabalho do Judiciário. Embora ela não elimine por completo os problemas, pode contribuir para a agilização e eficiência da tramitação dos processos, reduzindo custos e promovendo maior acessibilidade à Justiça.

2.5. A TRANSFORMAÇÃO DIGITAL DO PODER JUDICIÁRIO

A transformação digital (TD) no Judiciário brasileiro tem sido impulsionada pela necessidade de modernização e eficiência diante do alto volume de processos e da limitação de recursos humanos. Inserido no contexto da Revolução 4.0, que integra os universos físico, digital e biológico (Schwab, 2016), o Poder Judiciário tem buscado incorporar tecnologias

emergentes, como a IAG, para aprimorar a prestação jurisdicional. O contraste entre a posição modesta do Brasil no ranking global de competitividade digital (51ª entre 64 países em 2021) e a rápida adoção de tecnologias avançadas no judiciário, destacado por Toledo e Mendonça (2023), sugere uma implementação tecnológica que pode estar desalinhada com a maturidade do ecossistema digital nacional.

Os dados do Painel de Projetos de IA do CNJ 2024 confirmam essa implementação acelerada, mas também revelam questões preocupantes sobre a qualidade da governança. Enquanto houve um crescimento de 26% no número de projetos entre 2022 e 2023, apenas 26% dos projetos (37 de 140) estão devidamente registrados e em conformidade com as diretrizes do Sinapses (CNJ, 2024). Isso indica que a velocidade de implementação tem superado a capacidade de supervisão e controle institucional.

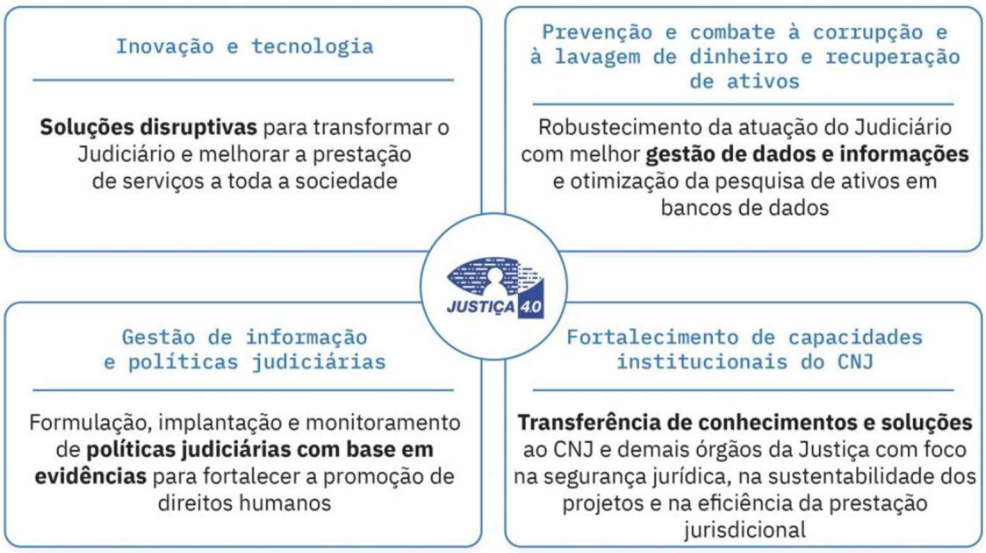
A pandemia da COVID-19, ao acelerar a digitalização das rotinas judiciais, intensificou também o uso de IA sem o correspondente desenvolvimento de salvaguardas éticas. Como apontam Bordoni e Tonet (2020), este processo foi impulsionado mais pela necessidade emergencial de manter o funcionamento do sistema durante o isolamento social do que por uma estratégia planejada de transformação digital eticamente orientada.

Este cenário de aceleração digital sem o devido acompanhamento ético manifesta-se em indicadores alarmantes identificados por Souza *et al*. (2024): questões relacionadas à privacidade e proteção de dados foram citadas em apenas 9 dos 45 estudos analisados, e a implementação e treinamento inadequados dos algoritmos em apenas 6, sugerindo uma subvalorização desses aspectos críticos na literatura e, possivelmente, na prática institucional.

2.5.1. Programa Justiça 4.0 - A Revolução Judicial em Curso

Diante desse cenário, o CNJ instituiu, em janeiro de 2021, o Programa Justiça 4.0, uma iniciativa desenvolvida em parceria com o Conselho da Justiça Federal (CJF), o Programa das Nações Unidas para o Desenvolvimento (PNUD) e outras instituições estratégicas, como o Tribunal Superior Eleitoral (TSE) e o Superior Tribunal de Justiça (STJ) (CNJ, 2024b). O programa está estruturado em quatro eixos estratégicos, como aponta a figura 11: (a) inovação e tecnologia; (b) gestão da informação e políticas judiciárias baseadas em evidências; (c) prevenção e combate à corrupção e à lavagem de dinheiro; e (d) fortalecimento das capacidades institucionais (Ferraz; Caracas; Baggio, 2022).

Figura 11: Os quatro eixos do Programa Justiça 4.0



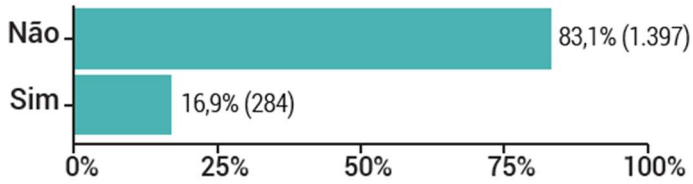
Fonte: CNJ (2024c)

Entre as soluções tecnológicas implementadas pelo Justiça 4.0 destacam-se a Plataforma Digital do Poder Judiciário (PDPJ-Br), a Plataforma Sinapses, a Plataforma Codex, o Juízo 100% Digital, o Balcão Virtual e os Núcleos de Justiça 4.0 (CNJ, 2024c). A PDPJ-Br, instituída pela Resolução n. 335/2020 do CNJ, busca transformar a tramitação processual em um sistema multisserviço adaptável às demandas de cada tribunal (Salomão *et al.,* 2023). Já o Juízo 100% Digital e o Balcão Virtual eliminaram a necessidade de comparecimento físico dos jurisdicionados aos fóruns, permitindo que audiências, julgamentos e atos processuais sejam realizados exclusivamente em ambiente digital, otimizando o acesso à Justiça e reduzindo custos operacionais (CNJ, 2024a; CNJ, 2024c).

O impacto dessas inovações tem sido expressivo. Em um estudo sobre o Tribunal de Justiça do Mato Grosso verificou-se que o tempo médio de duração dos processos de execução fiscal caiu de 2.450 dias para apenas 100 dias após a implementação dos Núcleos de Justiça 4.0 (Ferraz; Caracas; Baggio, 2022). Tais resultados evidenciam o potencial transformador da tecnologia na eficiência da prestação jurisdicional.

Contudo, os dados do relatório CNJ 2024 sobre IAG revelam que o sucesso quantitativo do programa não tem sido acompanhado de adequada governança ética. A figura 12 demonstra o resultado da pesquisa no qual 83,1% dos servidores e assessores que utilizaram IAG em atividades profissionais não informaram sobre esse uso, evidenciando uma lacuna crítica na transparência institucional (CNJ, 2024e). Este dado é particularmente preocupante quando se considera que a transparência é um dos princípios fundamentais da administração pública e um requisito essencial para a auditabilidade e controle dos processos judiciais.

Figura 12: Percentual de servidores(as) e/ou assessores(as) que informaram ter utilizado a IAG como ferramenta de apoio no desenvolvimento das atividades profissionais



Fonte: CNJ (2024e, p. 72)

O Programa Justiça 4.0, sem dúvida, representa um marco na transformação digital do judiciário brasileiro. Contudo, uma análise crítica de sua estruturação revela um desequilíbrio entre a ênfase em eficiência operacional e a atenção dedicada às questões éticas e de transparência. Dos quatro eixos estratégicos do programa - (a) inovação e tecnologia; (b) gestão da informação e políticas judiciárias baseadas em evidências; (c) prevenção e combate à corrupção e à lavagem de dinheiro; e (d) fortalecimento das capacidades institucionais (Ferraz; Caracas; Baggio, 2022) - nenhum aborda explicitamente questões relacionadas à ética algorítmica, transparência das decisões automatizadas ou proteção contra vieses discriminatórios.

2.5.2. Plataforma Sinapses: O Núcleo da Inteligência Artificial no Judiciário

Um dos pilares do Justiça 4.0 é a Plataforma Sinapses, instituída pela Resolução CNJ n. 332/2020, que define diretrizes para o armazenamento, treinamento supervisionado, controle de versionamento, distribuição e auditoria de modelos de IA aplicados à Justiça. A Sinapses foi inicialmente desenvolvida pelo Tribunal de Justiça de Rondônia (TJRO) em 2017 e, em 2018, começou a ser adaptada para uso nacional em parceria com o CNJ (CNJ, 2020). Em 2020, a plataforma foi reconhecida como vencedora do Prêmio de Inovação na categoria Prestação de Serviços da Expojud.

A Plataforma Sinapses viabiliza o compartilhamento de projetos de IA entre os tribunais, promovendo escalabilidade e padronização. Sua governança distribui a responsabilidade pelos modelos e *datasets* a cada órgão do Judiciário, enquanto o CNJ, por meio de seu Departamento de Tecnologia da Informação, mantém a infraestrutura operacional (CNJ, 2020). Essa abordagem descentralizada, mas coordenada, tem permitido maximizar a eficiência dos investimentos em IA e potencializar a adoção de soluções inovadoras no setor.

Entretanto, os dados do Painel de Projetos de IA 2024 revelam uma discrepância alarmante entre a proliferação de projetos de IA e sua adequada supervisão pela Plataforma Sinapses. Dos 140 projetos mapeados, apenas 37 estão registrados e em conformidade na plataforma - uma taxa de apenas 26% de aderência às diretrizes de governança estabelecidas pelo CNJ (CNJ, 2024d). Esta baixa taxa de *compliance* sugere que a maioria dos tribunais está desenvolvendo e implementando soluções de IA sem a devida supervisão institucional, o que representa riscos significativos para a uniformidade, qualidade e transparência dos sistemas.

Um aspecto crítico, raramente discutido nas avaliações institucionais da Plataforma, refere-se à qualidade e representatividade dos dados utilizados no treinamento dos modelos. Como apontam Maranhão, Junquilho e Tasso (2023), a ausência de protocolos rigorosos para avaliação e documentação dos conjuntos de dados utilizados compromete a capacidade de identificar e mitigar vieses sistemáticos. Esta lacuna é particularmente grave considerando a conhecida sub-representação de determinados grupos sociais nos dados do judiciário brasileiro, como evidenciado no relatório "Justiça em Números" (CNJ, 2024a).

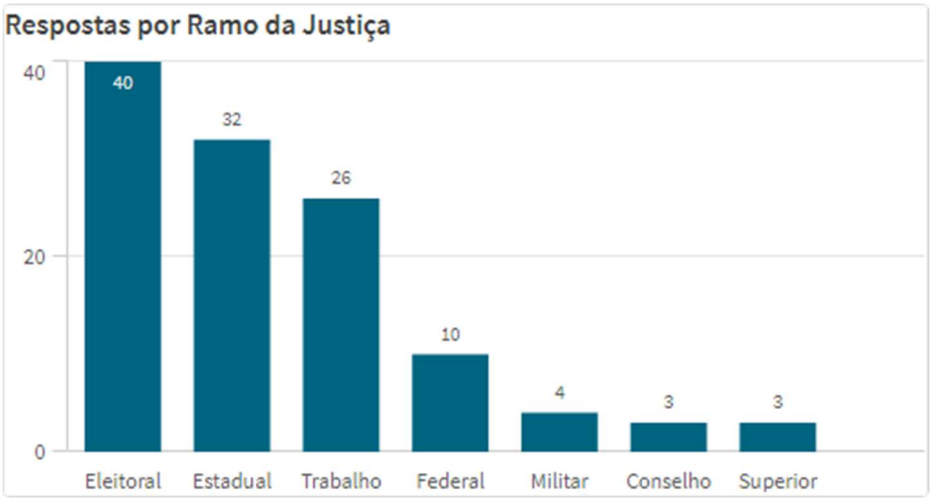
Outra dimensão insuficientemente tratada na governança da Plataforma Sinapses é a questão da responsabilidade por decisões automatizadas. A estrutura descentralizada, onde cada órgão do judiciário é responsável por seus modelos e *datasets*, pode diluir a responsabilidade pelos resultados, criando zonas cinzentas de *accountability*. Este problema é agravado pela natureza técnica e complexa dos sistemas de IA, que dificulta a supervisão efetiva por parte dos órgãos de controle tradicionais do judiciário.

Hoch e Engelmann (2023) observam que a regulação da IA no judiciário brasileiro ainda carece de mecanismos efetivos para responsabilização por danos causados por decisões algorítmicas equivocadas. Esta lacuna normativa representa um risco significativo para a proteção de direitos fundamentais, principalmente em um contexto de rápida expansão do uso dessas tecnologias em procedimentos jurisdicionais e administrativos.

2.5.3. Governança e Transparência Algorítmica

A governança e transparência algorítmica emergem como dimensões fundamentais para uma implementação ética da IA no judiciário brasileiro, especialmente à luz dos dados revelados pelo Painel de Pesquisa sobre IA 2024 do CNJ. Este levantamento, que identificou 98 projetos ativos de IA no Poder Judiciário nacional, evidencia tanto o crescimento exponencial dessas iniciativas quanto a urgência de mecanismos efetivos de controle e transparência. Como mostra a figura 13, a alta demanda de iniciativas na Justiça Estadual demonstra o protagonismo desta esfera no uso de IA.

Figura 13: Projetos de IA em 2024, por ramo de justiça



Fonte: Painel de pesquisa sobre Inteligência Artificial – 2024 (CNJ, 2024d)

A Resolução CNJ n. 332/2020 estabelece diretrizes para o desenvolvimento e uso de IA, abordando princípios fundamentais como não discriminação, publicidade, explicabilidade, controle do usuário e prestação de contas (CNJ, 2020). O normativo determina que os modelos de IA sejam auditáveis e que suas decisões possam ser interpretadas e questionadas, evitando o uso de sistemas opacos que comprometam a previsibilidade e a equidade das decisões judiciais. Contudo, os achados do Painel de Pesquisa 2024 revelam uma implementação heterogênea desses princípios, com variações significativas entre os tribunais quanto ao nível de documentação, auditoria e transparência dos sistemas implementados.

A análise do Tribunal de Justiça de Pernambuco (TJPE), conforme evidenciado em seu Plano de Tecnologia da Informação e Comunicação 2021-2026 (PTD), ilustra essa realidade de forma emblemática. O tribunal tem investido em iniciativas de modernização tecnológica e automação de processos, incluindo o desenvolvimento de soluções baseadas em IA para otimização de atividades administrativas e jurisdicionais. No entanto, a documentação disponível sugere uma ênfase predominante nos aspectos técnicos e operacionais, com menor detalhamento sobre os mecanismos de governança algorítmica e transparência das decisões automatizadas.

Os dados utilizados no treinamento dos algoritmos devem ser provenientes de fontes seguras e auditáveis, preferencialmente governamentais, garantindo rastreabilidade e conformidade com a Lei Geral de Proteção de Dados (LGPD) (Trento, 2021). Como medida de governança, os modelos devem passar por homologação para evitar vieses algorítmicos que possam comprometer a imparcialidade da justiça (CNJ, 2020). O Painel de Pesquisa 2024

indica, porém, que apenas 32% dos projetos de IA em funcionamento no judiciário possuem protocolos formais de auditoria algorítmica, sugerindo uma lacuna significativa na implementação efetiva desses princípios.

Maranhão, Junquilho e Tasso (2023) identificam uma lacuna entre os princípios declarados e os mecanismos concretos para sua efetivação, especialmente no que se refere à explicabilidade dos algoritmos. Os autores propõem um modelo de governança que exigiria a divulgação pública de documentação técnica detalhada sobre os sistemas de IA, permitindo auditoria externa – proposta que encontra resistência institucional sob alegações de segurança e propriedade intelectual. Esta tensão torna-se particularmente evidente quando contrastada com os dados do Painel CNJ 2024, que mostram que apenas 18% dos tribunais disponibilizam informações técnicas detalhadas sobre seus sistemas de IA para consulta pública.

A experiência internacional demonstra que a transparência algorítmica é não apenas desejável eticamente, mas também tecnicamente viável. Como apontam Fornasier, Silva e Schwede (2023), jurisdições como a União Europeia têm avançado em *frameworks* regulatórios que exigem explicabilidade e auditabilidade de sistemas automatizados de tomada de decisão, sem comprometer sua eficácia operacional. A limitada adesão a esses padrões no judiciário brasileiro sugere mais uma escolha institucional que uma impossibilidade técnica.

Um aspecto particularmente problemático da governança algorítmica no contexto brasileiro é a limitada participação da sociedade civil e de comunidades potencialmente afetadas no desenvolvimento e avaliação dos sistemas. Conforme destacado por Cambi e Amaral (2023), a ausência de mecanismos participativos na governança da IA judicial compromete não apenas a identificação precoce de vieses, mas também a legitimidade social das soluções implementadas. Este déficit democrático na governança tecnológica contrasta com os princípios constitucionais de publicidade e participação que deveriam orientar a administração pública brasileira. O Painel CNJ 2024 confirma essa preocupação ao revelar que menos de 15% dos projetos de IA no judiciário incluem mecanismos formais de consulta pública ou participação da sociedade civil em sua concepção e implementação.

2.6. A EXPERIÊNCIA DO TJPE NA TRANSFORMAÇÃO DIGITAL E NO USO DA INTELIGÊNCIA ARTIFICIAL

O Tribunal de Justiça de Pernambuco (TJPE) tem conquistado posição de destaque no cenário judiciário nacional, figurando consistentemente entre os primeiros tribunais estaduais do Brasil em diversos indicadores de desempenho. Essa posição evidencia-se nos resultados divulgados pelo Conselho Nacional de Justiça (CNJ) através do Relatório Justiça em Números

de 2024, principal instrumento de avaliação do Poder Judiciário brasileiro que coloca o TJPE "entre os primeiros no ranking dos 27 Tribunais estaduais do país" (CNJ, 2024a).

A consolidação do uso da inteligência artificial generativa (IAG) no setor público depende não apenas de diretrizes normativas e modelos de governança, mas também da existência de ambientes institucionais que viabilizem sua experimentação segura e estruturada. Nesse contexto, o TJPE constitui um exemplo representativo de maturidade digital e protagonismo nacional, reunindo características técnicas e operacionais que favorecem a implementação de soluções inovadoras baseadas em IA.

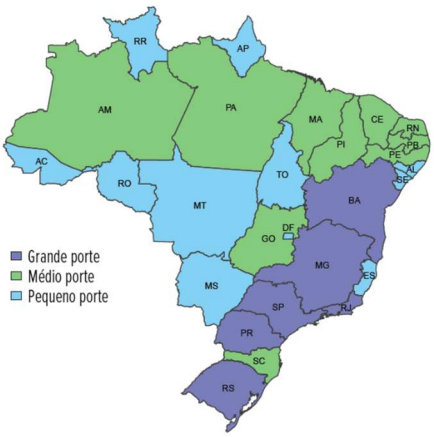
Conforme delineado em seu Plano de Transformação Digital (PTD 2021–2026), o TJPE estabeleceu metas ambiciosas em cinco eixos estratégicos: infraestrutura, interoperabilidade, gestão da informação, transformação dos serviços e governança e gestão (TJPE, 2022). Para atender a essas diretrizes, o tribunal investiu na modernização do parque computacional, com foco na substituição de mais de 8.800 estações de trabalho obsoletas, na ampliação da conectividade das unidades judiciárias e na adoção de serviços em nuvem para garantir escalabilidade e segurança da informação.

Além da infraestrutura tecnológica, destaca-se o esforço do tribunal em adotar soluções baseadas em inteligência artificial, incluindo automações por RPA[4], ferramentas de triagem processual e modelos preditivos aplicados ao acervo judicial. O uso dessas tecnologias está alinhado à Plataforma Digital do Poder Judiciário (PDPJ-Br) e ao Programa Justiça 4.0, impulsionando melhorias na produtividade, na celeridade e na qualidade da prestação jurisdicional.

Dados extraídos do Painel Nacional de IA do CNJ (2024) corroboram esse protagonismo: o TJPE está classificado como tribunal de médio porte, conforme demonstrado na figura 14, e se encontra entre os tribunais estaduais com maior número de projetos ativos envolvendo IA, principalmente nos eixos de automação de tarefas e auxílio à tomada de decisão. Esse posicionamento confere ao tribunal não apenas visibilidade institucional, mas também representatividade suficiente para que as experiências locais sirvam de base para replicações em outros órgãos do Judiciário, sobretudo diante da sua realidade processual heterogênea e do seu papel de articulação regional no Nordeste.

[4] Automação Robótica de Processos (RPA) refere-se à tecnologia que utiliza robôs de software (bots) para automatizar tarefas repetitivas e estruturadas, emulando ações humanas em sistemas digitais. Diferentemente da Inteligência Artificial Generativa, a RPA executa processos predefinidos sem capacidade de aprendizagem ou criação de conteúdo original, sendo particularmente eficaz na automação de atividades administrativas de alto volume no judiciário (IBM, 2024).

Figura 14: Distribuição territorial dos Tribunais de Justiça segundo o porte.



Fonte CNJ (2024ª, pág. 59)

O posicionamento do TJPE no cenário nacional de transformação digital evidencia características que justificam sua seleção como objeto empírico desta pesquisa. Entre os tribunais estaduais brasileiros, o TJPE destaca-se não apenas pelos indicadores de desempenho reconhecidos pelo CNJ, mas também por sua maturidade em projetos de inteligência artificial e participação ativa na Plataforma Sinapses. O Quadro 4 apresenta um panorama comparativo que contextualiza o perfil de transformação digital dos principais tribunais de justiça do país.

Quadro 4: Perfil de Transformação Digital dos Principais Tribunais de Justiça

Tribunal	Porte	Projetos IA*	Sinapses	Características Distintivas
TJRS	**Grande**	**12**	✓	• Maior número de projetos de IA (liderança nacional) • Pioneiro na Plataforma Sinapses (2017) • Centro de excelência em inovação judicial
TJCE	**Médio**	**7**	✓	• Alto índice de projetos/porte (eficiência) • Forte adesão ao Programa Justiça 4.0 • Referência em gestão baseada em dados
TJSP	**Grande**	**6**	✓	• Maior tribunal do país em volume processual • Projetos focados em automação em larga escala • Integração com ecossistema tecnológico paulista
TJPE	**Médio**	**4**	✓	• **Ranking entre os primeiros TJs do país (CNJ)** • **PTD 2021-2026 com 5 eixos estratégicos** • **Modernização de 8.800 estações de trabalho**
TJMG	**Grande**	**5**	✓	• Pioneiro em uso de ChatGPT (relatório CNJ) • Foco em IA para análise jurisprudencial • Parcerias com universidades locais
TJDFT	**Médio**	**3**	✓	• Uso de IAG mencionado no relatório CNJ • Projetos piloto em automação administrativa • Integração com iniciativas do Poder Executivo
Outros	**Variado**	**1-3**	**Parcial**	• Estágios iniciais de implementação • Foco em digitalização básica • Projetos isolados sem integração sistêmica

Legenda e Critérios:

Grande: >10M processos anuais
Médio: 3-10M processos anuais
Pequeno: <3M processos anuais

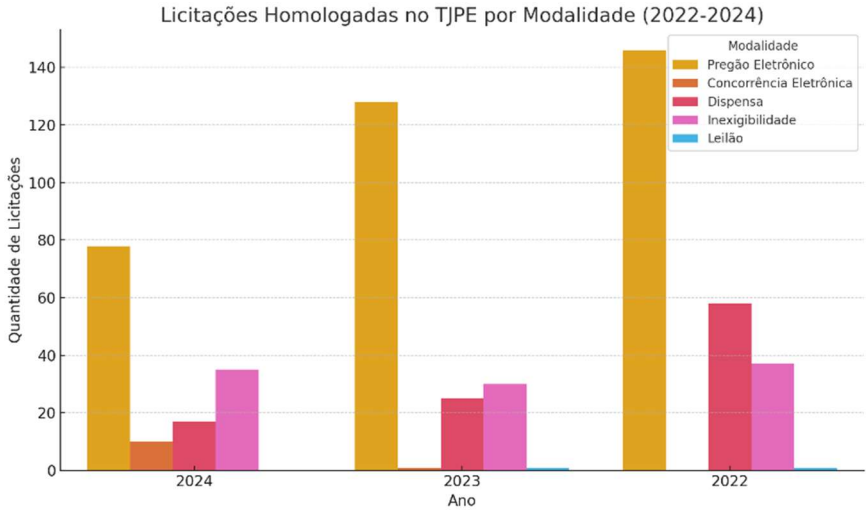
✓ **Sinapses:** Participação ativa
Parcial **Parcial:** Adesão limitada

Notas:
*Projetos de IA: Baseado no Painel CNJ 2024 (140 projetos mapeados)
**TJPE em destaque: Dados confirmados pela instituição e relatórios CNJ

Fonte: Baseado em CNJ (2024a, 2024d, 2024e) e TJPE (2022)

No que diz respeito às contratações públicas, o TJPE também apresenta um perfil propício à adoção de IAG para qualificação de artefatos licitatórios. Entre 2022 e 2024, o tribunal homologou 565 processos licitatórios, sendo 100 por dispensa e 102 por inexigibilidade – o que representa 35,7% do total, conforme ilustrado na figura 15. A expressiva presença de contratações diretas reforça a relevância de iniciativas voltadas à padronização e ao uso de IA na elaboração de documentos como termos de referência, minutas e relatórios de homologação

Figura 15: Licitações homologadas no TJPE por modalidade e ano (2022–2024)



Fonte: Baseado em dados do TJPE (2025)

Desse modo, o TJPE não apenas reflete o movimento nacional de modernização do Judiciário, como também desempenha um papel estratégico na experimentação e disseminação de práticas baseadas em IAG. Sua estrutura organizacional, ambiente digitalizado e cultura institucional voltada à inovação justificam plenamente sua escolha como unidade empírica deste estudo de caso.

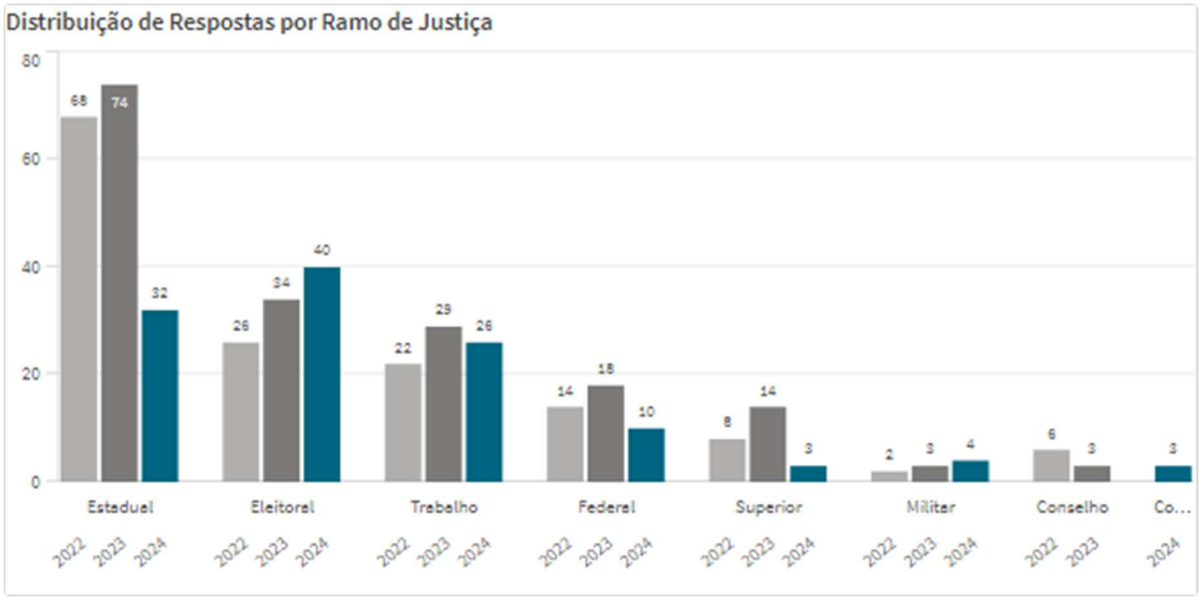
2.7. DESAFIOS E PERSPECTIVAS ACERCA DA ÉTICA E DA TRANSPARÊNCIA

A implementação de tecnologias de Inteligência Artificial no Poder Judiciário Pernambucano ocorre em um cenário nacional de crescimento exponencial e desafios emergentes. As evidências mais recentes apresentadas pelo CNJ através da Pesquisa de Uso de IA no Poder Judiciário 2023 e do relatório "O uso da Inteligência Artificial Generativa no Poder Judiciário Brasileiro" revelam uma realidade complexa que demanda análise criteriosa e ações estruturantes por parte do TJPE (CNJ, 2024e).

Nessa linha, o panorama nacional demonstra um crescimento significativo na adoção de soluções baseadas em IA, como evidenciado na figura 16, que ilustra a distribuição dessas

tecnologias no país. O aumento quantitativo de iniciativas, embora positivo sob a perspectiva da modernização, suscita preocupações sobre a qualidade da implementação e seus impactos nos direitos fundamentais dos jurisdicionados.

Figura 16: Total de projetos de IA no judiciário brasileiro entre 2022 e 2024



Fonte: CNJ (2025)

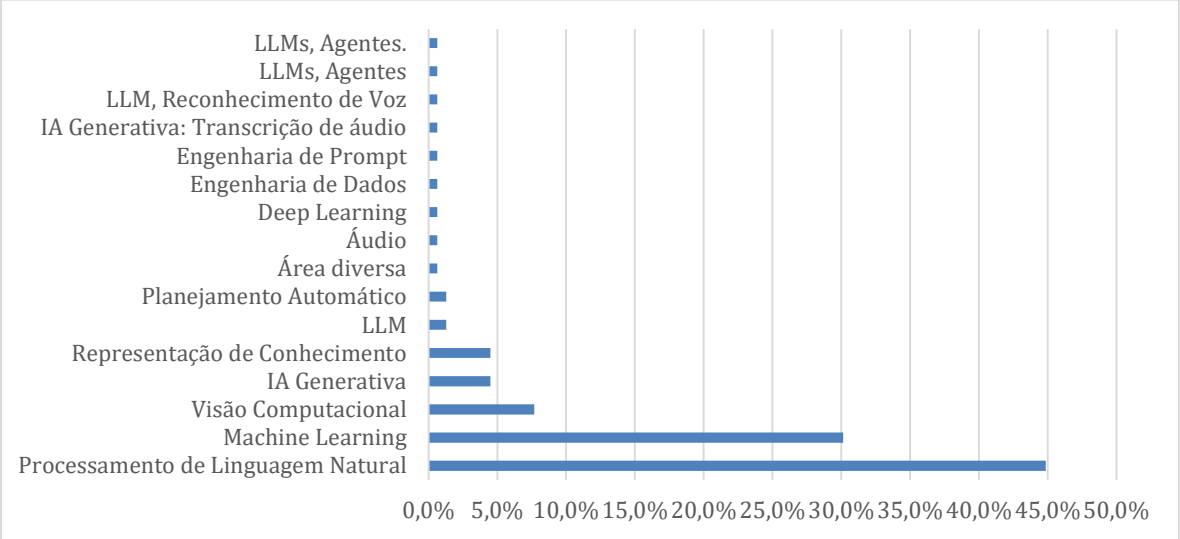
As preocupações éticas identificadas por Souza *et al*. (2024) no mapeamento sistemático da literatura ganham contornos mais precisos quando confrontadas com dados empíricos recentes. Quase metade do total de magistrados (49,4%) e servidores (49,5%) afirmou que utiliza a IA generativa (CNJ, 2024e), revelando uma adoção massiva que, paradoxalmente, ocorre de forma não transparente. Cerca de 85% dos usuários não informa sobre esse uso em suas atividades profissionais (CNJ, 2024e), configurando um fenômeno de "*shadow IT*" que compromete a governança institucional e a auditabilidade dos processos.

A figura 17 expõe a análise das subáreas de IA utilizadas no Poder Judiciário em 2024, revelando a predominância de tecnologias de processamento de linguagem natural (PLN) e aprendizado de máquina supervisionado, seguidas por sistemas de classificação automatizada e análise preditiva. Essa distribuição indica que as principais aplicações se concentram em automatização de tarefas repetitivas e apoio à decisão, áreas que, embora promissoras para ganhos de eficiência, apresentam riscos significativos quando implementadas sem adequada supervisão (CNJ, 2025).

As principais barreiras para uma implementação eticamente orientada da IA no judiciário, conforme apontado por Toledo e Mendonça (2023), são a capacitação insuficiente

dos servidores (30,8%), as deficiências na digitalização e tratamento adequado dos dados (28,8%) e as limitações na infraestrutura tecnológica (25%). Estes obstáculos, embora frequentemente tratados como meramente técnicos ou operacionais, têm profundas implicações éticas, na medida em que comprometem a capacidade do sistema de garantir equidade, precisão e responsabilidade em suas operações automatizadas.

Figura 17: Subáreas de IA utilizadas

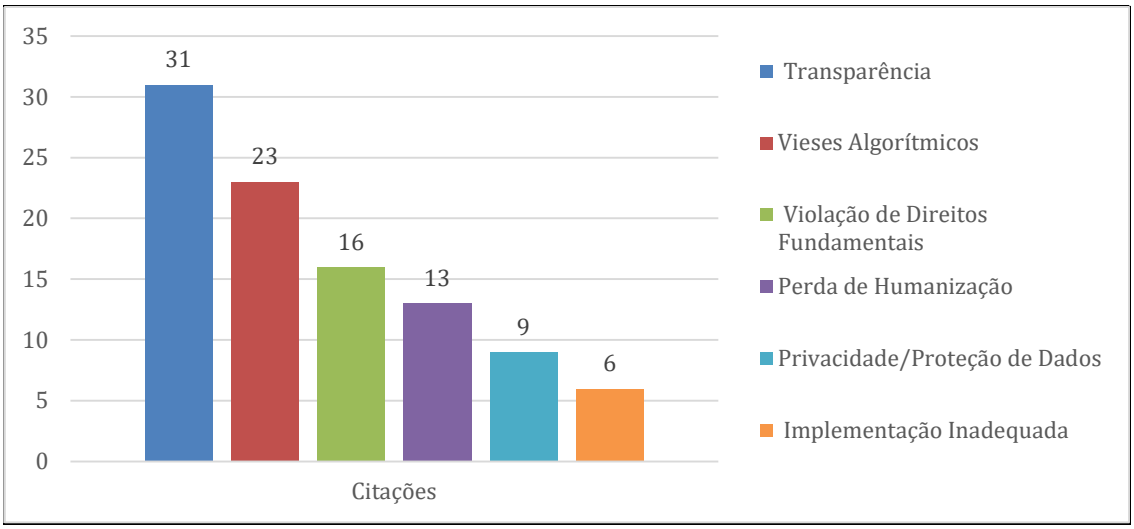


Fonte: (CNJ, 2025)

A figura 18 exibe as principais preocupações no uso de IA no judiciário, segundo Souza et al. (2024), que identificam uma preocupação crescente na literatura com a perda da humanização nas decisões judiciais (13 citações em 45 artigos), evidenciando tensões entre a automação e valores como a empatia e a consideração contextualizada dos casos – dimensões tradicionalmente associadas à justiça substantiva. Como enfatizam Casimiro e Gouveia (2023), a substituição do julgamento humano por sistemas automatizados pode comprometer a capacidade do judiciário de responder adequadamente à complexidade social e moral das questões que lhe são apresentadas.

Outro desafio refere-se à prestação de contas por decisões automatizadas. Fornasier (2023) destaca que o "direito à explicação", fundamental em sistemas judiciais democráticos, enfrenta obstáculos técnicos significativos quando decisões são tomadas por modelos complexos de aprendizado profundo, cuja lógica interna pode ser opaca mesmo para seus desenvolvedores. Esta tensão entre eficiência técnica e transparência democrática permanece insuficientemente endereçada nas políticas de implementação de IA no judiciário brasileiro.

Figura 18: Principais desafios no uso de IA no judiciário



Fonte: Baseado em Souza *et al.* (2024)

A discriminação algorítmica emerge como uma preocupação central, especialmente em um país marcado por profundas desigualdades sociais e raciais. Cambi e Amaral (2023) argumentam que, sem mecanismos robustos de auditoria e mitigação de vieses, sistemas de IA no judiciário podem não apenas reproduzir, mas amplificar padrões históricos de discriminação, comprometendo o princípio constitucional da igualdade. A escassez de estudos empíricos sobre vieses algorítmicos no contexto judicial brasileiro representa uma lacuna que dificulta o desenvolvimento de soluções efetivas para este problema.

Prado, Münch e Villarroel (2022) enfatizam a importância da formação ética dos magistrados e servidores para o uso responsável da IA, destacando que a tecnologia, por si só, não garante decisões justas – estas dependem fundamentalmente do enquadramento ético e dos valores que orientam sua implementação e uso. A capacitação humanística, e não apenas técnica, dos profissionais envolvidos, emerge como um componente essencial para uma transformação digital que preserve os princípios fundamentais do Estado Democrático de Direito.

A Resolução nº 615/2025 do CNJ, que entrará em vigor a partir de 11/09/2025 (CNJ, 2025a) atualiza e amplia as diretrizes estabelecidas pela Resolução nº 332/2020, visando regulamentar o uso da IA no Poder Judiciário de forma mais abrangente e alinhada às evoluções tecnológicas recentes. O quadro 5 destaca as principais mudanças da nova resolução.

A Resolução nº 615/2025 representa um avanço significativo na regulamentação do uso de IA no Poder Judiciário brasileiro, ao estabelecer normas mais detalhadas e rigorosas que visam garantir a ética, a transparência, a segurança e a proteção dos direitos fundamentais no desenvolvimento e aplicação dessas tecnologias. Esta evolução normativa reforça a necessidade

de que o TJPE se antecipe às exigências regulamentares, desenvolvendo, desde o início, práticas que excedam os padrões mínimos de *compliance*.

Para o TJPE, esses achados representam tanto oportunidades quanto advertências. A posição de destaque da instituição no cenário nacional, conforme evidenciado em seu PTD 2021-2026, oferece condições favoráveis para implementar práticas de vanguarda em governança algorítmica. Contudo, a experiência nacional demonstra que o entusiasmo tecnológico deve ser temperado por rigorosos protocolos éticos e de transparência.

A experiência acumulada pelo Judiciário brasileiro evidencia que o sucesso na implementação de IA depende não apenas da sofisticação tecnológica, mas fundamentalmente da capacidade institucional de estabelecer marcos éticos claros, processos de governança efetivos e cultura organizacional orientada pela transparência e responsabilidade, alinhado às demandas de uma sociedade que busca celeridade e eficiência na prestação jurisdicional.

Para o TJPE, esse aprendizado deve orientar todas as iniciativas relacionadas à aplicação de IAG em seus processos administrativos, assegurando que a modernização tecnológica fortaleça, ao invés de comprometer, os princípios fundamentais que regem a administração pública brasileira. A tecnologia deve servir como um instrumento para a realização mais plena da justiça, e não como um fim em si mesma. No entanto, para que essa tecnologia alcance seu potencial máximo, é fundamental enfrentar desafios relacionados à capacitação, infraestrutura, governança e monitoramento de impactos.

Quadro 5: Resolução CNJ nº 332/2020 × Resolução CNJ nº 615/2025

Tema	Resolução nº 332/2020	Resolução nº 615/2025
Objetivo Geral	Estabelecer diretrizes éticas, de transparência e governança para o uso de IA no Judiciário.	Regulamentar o desenvolvimento, governança, auditoria, monitoramento e uso responsável de soluções de IA, com foco na inovação tecnológica e eficiência dos serviços judiciários, assegurando a observância dos direitos fundamentais.
Supervisão Humana	Recomenda supervisão humana nas decisões assistidas por IA.	Exige supervisão humana efetiva, periódica e adequada em todas as etapas do ciclo de vida da IA, proibindo decisões autônomas sem possibilidade de revisão humana.
Classificação de Riscos	Não prevê classificação formal de riscos.	Introduz classificação de soluções de IA em três categorias: baixo risco, alto risco e risco excessivo, com exigências proporcionais de auditoria e monitoramento conforme o impacto potencial nos direitos fundamentais.
Vedações Específicas	Não detalha proibições específicas quanto ao uso de IA.	Proíbe expressamente o uso de IA para: decisões autônomas sem revisão humana; valoração de traços de personalidade ou comportamentos para prever crimes ou reincidência; classificação de pessoas com base em características pessoais; reconhecimento biométrico de emoções.

Governança e Auditoria	Estabelece diretrizes gerais de governança e qualidade.	Define processos internos obrigatórios para garantir segurança, transparência e auditabilidade dos sistemas de IA, incluindo a publicação de relatórios detalhados sobre funcionamento, finalidades, dados utilizados e mecanismos de supervisão.
Proteção de Dados	Determina que os dados utilizados devem ser provenientes de fontes seguras e protegidos contra riscos de destruição, modificação e transmissões não autorizadas.	Incorpora os princípios de "*privacy by design*" e "*privacy by default*", exigindo anonimização dos dados desde a concepção dos sistemas, avaliações de impacto obrigatórias e conformidade com a Lei Geral de Proteção de Dados (LGPD).
Transparência e Prestação de Contas	Recomenda a explicação dos passos que conduziram ao resultado das decisões assistidas por IA.	Exige a publicação de indicadores claros e relatórios públicos em linguagem simples, garantindo que os jurisdicionados tenham ciência do uso de IA e compreendam como essas tecnologias estão sendo aplicadas nos processos judiciais.
Participação Institucional	Prevê que os órgãos do Judiciário informem ao CNJ sobre a pesquisa, desenvolvimento, implantação e uso de IA.	Cria o Comitê Nacional de Inteligência Artificial do Judiciário, com composição plural, incluindo representantes do CNJ, magistratura, OAB, Ministério Público, Defensoria Pública e sociedade civil, responsável por auxiliar na implementação, cumprimento e supervisão da aplicação da resolução.
Capacitação e Letramento Digital	Não aborda especificamente a capacitação dos usuários.	Estabelece a necessidade de capacitação contínua dos usuários das soluções de IA, promovendo o letramento digital como condição prévia para o uso seguro e legítimo da tecnologia no Judiciário.
IA Generativa	Não abordado	Capítulo específico (VI) - arts. 19-21

Fonte: Elaborado pela autora (2025)

REFERÊNCIAS

ALENCAR, A. C. de; BLIACHERIENE, A. C.; ARAÚJO, L. V. de. Um modelo de e-marketplace para compras públicas eficazes com o uso de inteligência artificial generativa. **Revista de Direitos e Garantias Fundamentais**, v. 25, n. 3, p. 139-168, 2024. Disponível em: https://sisbib.emnuvens.com.br/direitosegarantias/article/view/2496. Acesso em: 20 mai. 2025.

BRASIL. Decreto nº 12.198, de 24 de setembro de 2024. Institui a Estratégia de Governo Digital para o período de 2024 a 2027, no âmbito dos órgãos e das entidades da administração pública federal. **Diário Oficial da União**, Brasília, DF, 24 set. 2024. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2024/decreto/d12198.htm. Acesso em 20 mai. 2025.

BORDONI, J. A.; TONET, L. Inovação e Tecnologia no Judiciário. **Themis: Revista da Esmec**, v. 18, n. 2, p. 151-170, 2020. Acesso em: 02 out. 2024. Disponível em: https://revistathemis.tjce.jus.br/index.php/THEMIS/article/view/792.

BRITO, T. S.; FERNANDES, R. S. Inteligência Artificial e a Crise do Poder Judiciário: Linhas Introdutórias sobre a Experiência Norte-Americana, Brasileira e sua Aplicação no Direito Brasileiro. **Revista Acadêmica da Faculdade de Direito do Recife**, v. 91, n. 2, p. 84-107, 2020. Acesso em: 02 out. 2024. Disponível em: https://periodicos.ufpe.br/revistas/index.php/ACADEMICA/article/view/247757.

BRITO, T. S.; FERNANDES, R. S. Crise do Judiciário e eficiência na resolução de conflitos: uma abordagem contemporânea. **Revista de Direito Processual**, v. 3, n. 1, p. 56-78, 2022.

CAMBI, Eduardo Augusto Salomão; AMARAL, Maria Eduarda Toledo Pennacchi Tibiriçá. Inteligência artificial no poder judiciário, discriminação algorítmica e direitos humanos-fundamentais. **Suprema–Revista de Estudos Constitucionais,** 3(2), 189-218, 2023. Acesso em: 27 mar. 2025. Disponível em: https://www.researchgate.net/publication/376492933_Inteligencia_artificial_no_Poder_Judiciario_discriminacao_algoritmica_e_direitos_humanos-fundamentais.

CASIMIRO, J. S. C. ; GOUVEIA, L. G. de. Processo judicial e decisão fundamentada. Atualmente no contexto brasileiro é possível a inteligência artificial de raciocínio jurídico aplicar o direito tal qual o juiz humano? **Revista Eletrônica de Direito Processual**, Rio de Janeiro, v. 24, n. 3, 2023. DOI: 10.12957/redp.2023.74005. Disponível em: https://www.e-publicacoes.uerj.br/redp/article/view/74005. Acesso em: 27 mar. 2025.

CHATGPT BRASIL. **ChatGPT vs Gemini vs Copilot**. 2024. Disponível em: https://chatgpt.com.br. Acesso em: 04 jun. 2025.

COMPUTER HOY**. ChatGPT vs Gemini vs Copilot:** ¿Cuál es el mejor chatbot de IA*?* 2024. Disponível em: https://computerhoy.20minutos.es. Acesso em: 04 jun. 2025.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). **Plataforma SINAPSES**. Brasília: CNJ, 2025. Disponível em: https://www.cnj.jus.br/sistemas/plataforma-sinapses/. Acesso em: 14 mar. 2025.

_____. **Resolução nº 332, de 21 de agosto de 2020**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. Brasília: CNJ, 2020. Disponível em: https://atos.cnj.jus.br/atos/detalhar/3429. Acesso em: 02 out. 2024.

_____. **Justiça em números 2023**. Brasília: CNJ, 2024a. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2024/05/justica-em-numeros-2024.pdf. Acesso em 02 out. 2024.

_____. **Justiça 4.0**. 2024b. Disponível em: https://www.cnj.jus.br/tecnologia-da-informacao-e-comunicacao/justica-4-0/. Acesso em: 03 out. 2024.

_____. **Cartilha Justiça 4.0**. Brasília: CNJ, 2024c. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2024/05/cartilha-justica-4-0-pt-2024.pdf. Acesso em: 03 out. 2024.

_____. **Painel de projetos de inteligência artificial no Judiciário – 2024.** Brasília: CNJ, 2024d. Acesso em: 30 mai. 2025. Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=51977be5-96d0-4362-98ff-ed3eb3337781.

_____. **O uso da inteligência artificial generativa no Poder Judiciário brasileiro** : relatório de pesquisa / Conselho Nacional de Justiça. – Brasília: CNJ, 2024e. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2024/09/cnj-relatorio-de-pesquisa-iag-pj.pdf. Acesso em 30 mai. 2025

_____. **Resolução nº 615, de 11 de março de 2025**. Estabelece diretrizes desenvolvimento, para utilização o e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário. Brasília: CNJ, 2025. Acesso em: 05 jun. 2025. Disponível em: https://atos.cnj.jus.br/files/original155530202503146 7d4517244566.pdf.

CONTINI, F. Unboxing Generative AI for the Legal Professions: Functions, Impacts and Governance. **International Journal for Court Administration**, v. 15, n. 2, p. 1-22, 2024. Disponível em: https://iacajournal.org/en/articles/10.36745/ijca.604. Acesso em: 28 set. 2024.

FERRAZ, T. S.; CARACAS, J. R.; BAGGIO, C. T. P. Programa Justiça 4.0: a perspectiva inovadora da prestação jurisdicional sob o enfoque da celeridade e da transparência. **Brazilian Journal of Development**, v. 8, n. 10, p. 67590-67610, 2022. Acesso em: 02 out. 2024. Disponível em: https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/53190.

FORNASIER, M. de O. Deep learning and the right to explanation: technological challenges to legality and due process of law. **Revista de Direito Brasileira**, Florianopolis, Brasil, v. 32, n. 12, p. 218–235, 2023. DOI: 10.26668/IndexLawJournals/2358-1352/2022.v32i12.7547. Acesso em: 27 mar. 2025. Disponível em: https://indexlaw.org/index.php/rdb/article/view/7547.

FORNASIER, M. de O.; SILVA, F. V. da; SCHWEDE, M. A. A utilização de ferramentas de inteligência artificial no judiciário brasileiro e a resolução 332/2020 do CNJ. **Revista Jurídica Cesumar - Mestrado**, *[S. l.]*, v. 23, n. 2, p. 275–288, 2023. DOI: 10.17765/2176-9184.2023v23n2.e10435. Acesso em: 27 mar. 2025. Disponível em: https://periodicos.unicesumar.edu.br/index.php/revjuridica/article/view/10435.

GROWTH SALES. **Comparativo ChatGPT4 vs Copilot vs GeminiPro**. 2024. Disponível em: https://growthsales.es. Acesso em: 04 jun. 2025.

HOCH, P. A.; ENGELMANN, W. Regulação da inteligência artificial no judiciário brasileiro e europeu. **Pensar-Revista de Ciências Jurídicas**, v. 28, n. 4, p. 1-18, 2023. Disponível em: https://ojs.unifor.br/rpen/article/view/14263. Acesso em: 29 fev. 2024.

IBM. **O que é automação robótica de processos (RPA)?** 2024. Disponível em: https://www.ibm.com/br-pt/topics/rpa. Acesso em: 30 mai. 2025.

KAPLAN, A.; HAENLEIN, M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. **Business Horizons**, v. 62, n. 1, p. 15-25, 2019. Acesso em: 31 mar. 2024. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S0007681318301393.

LIMBERGER, T.; GIANNAKOS, D. B. da S.; SZINVELSKI, M. M. Can Judges be Replaced by Machines? The Brazilian Case. **Mexican Law Review**, v. 14, n. 2, p. 53-81, 2022. Acesso em: 10 maio 2024. Disponível em: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-05782022000100053&lng=es&nrm=iso.

LIU, P.; YUAN, W.; FU, J.; JIANG, Z.; HAYASHI, H.; NEUBIG, G. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. **ACM Computing Surveys**, v. 55, n. 9, p. 1-35, 2023. Acesso em: 07 out. 2024. Disponível em: https://dl.acm.org/doi/abs/10.1145/3560815.

MARANHÃO, J. S.de A.; JUNQUILHO, T. A.; TASSO, F. A. Transparência sobre o emprego de inteligência artificial no judiciário: um modelo de governança. **Suprema: Revista de Estudos Constitucionais,** v. 3, p. [145]-187, 2023. Disponível em: https://legalgroundsinstitute.com/blog/transparencia-sobre-o-emprego-de-inteligencia-artificial-no-judiciario-um-modelo-de-governanca/. Acesso em: 28 fev 2024.

McCARTHY, J. **What is Artificial Intelligence?** Stanford University, 2007. Disponível em: http://jmc.stanford.edu/articles/whatisai/whatisai.pdf. Acesso em: 28 set. 2024.

PEREIRA, S. L. **Processamento de Linguagem Natural**. [S. l.: s. n.], [s.d.]. Disponível em: https://www.ime.usp.br/~slago/IA-pln.pdf. Acesso em: 08 mar. 2025.

PEREIRA, D. I.; BOTELHO, E. S. Agentes racionais na inteligência artificial aplicada ao direito. **Revista de Direito, Tecnologia e Sociedade**, v. 4, n. 2, p. 27-43, 2018.

PRADO, E. M. B.; MÜNCH, L. A. C.; VILLARROEL, M. A. C. U. "SOB CONTROLE DO USUÁRIO": FORMAÇÃO DOS JUÍZES BRASILEIROS PARA O USO ÉTICO DA IA NO JUDICIÁRIO. **Direito Público**, *[S. l.]*, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.6021. Disponível em: https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6021. Acesso em: 29 fev. 2024.

RAMOS, L. O.; CUNHA, L. G.; OLIVEIRA, F. L.; SAMPAIO, J. O. **Relatório ICJBrasil 2021**. FGV DIREITO SP, 2021. Disponível em: https://repositorio.fgv.br/server/api/core/bitstreams/82935cd1-3393-4262-80a6-e8e39570caf7/content. Acesso em: 02 out. 2024.

ROSSETTI, R.; GARCIA, K. Inteligência artificial generativa: questões jurídicas e éticas em torno do ChatGPT. **VirtuaJus**, Belo Horizonte, v. 8, n. 15, p. 253-264, 2023. Disponível em: https://periodicos.pucminas.br/index.php/virtuajus/article/download/30769/21848. Acesso em: 01 out. 2024.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence:** A Modern Approach. 3. ed. Harlow (UK): Pearson, 2013.

SAHOO, P.; SINGH, A. K.; SAHA, S.; JAIN, V.; MONDAL, S.; CHADHA, A. **A Systematic Survey of *Prompt* Engineering in Large Language Models**: Techniques and Applications. 2024. Disponível em: https://arxiv.org/abs/2402.07927. Acesso em: 04 out. 2024.

SALOMÃO, L. F.; TAUK, C. S. **Inteligência Artificial**: tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário brasileiro. 3. ed. Rio de Janeiro: FGV, 2023. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/relatorio_ia_3a_edicao_0.pdf. Acesso em: 02 out. 2024.

SCHWAB, K. **A quarta revolução industrial**. São Paulo: Edipro, 2016.

SHABBIR, J.; ANWER, T. Artificial Intelligence and its Role in Near Future. **Journal of Latex Class Files**, v. 14, n. 8, p. 1-11, 2015.

SHEN, Y.; KANG, K.; WANG, Y.; ZHANG, H.; WANG, W.; CHEN, W.; ZHANG, Y.; HAN, Z.; LI, Y.; GU, J.; YANG, W.; DING, Y.; ZOU, L.; ZOU, X.; SHOU, L.; ZHOU, M.; KIM, D. **Large language models: A Survey**. 2023. Disponível em: https://arxiv.org/abs/2303.18223. Acesso em: 01 out. 2024.

SIQUEIRA, O. N.; ZANFERDINI, F. A. M. Online Dispute Resolution e Inteligência Artificial: a influência tecnológica na resolução de conflitos. **Revista Eletrônica Direito e Sociedade-REDES**, v. 9, n. 2, p. 87-104, 2021. Acesso em: 02 out. 2024. Disponível em: https://revistas.unilasalle.edu.br/index.php/redes/article/view/7142.

SOUZA, L. B. dos S. B. de; ARAÚJO, A. C.; CORREIA NETO, J. da S.; MORAES, I. C. de. (2024). Preocupações associadas à utilização da inteligência artificial no judiciário brasileiro: um mapeamento sistemático da literatura. **International Journal of Scientific Management and Tourism**, *[S. l.]*, v. 10, n. 4, p. e1010, 2024. DOI: 10.55905/ijsmtv10n4-003. Disponível em: https://ojs.scientificmanagementjournal.com/ojs/index.php/smj/article/view/1010. Acesso em: 24 jul. 2024.

STONE, P.; BROOKS, R.; BRYNJOLFSSON, E.; CALO, R.; ETZIONI, O.; HAGER, G.; HIRSCHBERG, J.; KALYANAKRISHNAN, S.; KAMAR, E.; KRAUS, S.; LEYTON-BROWN, K.; PARKES, D.; PRESS, W.; SAXENIAN, A.; SHAH, J.; TAMBE, M.; TELLER, A. **Artificial Intelligence and Life in 2030**: One Hundred Year Study on Artificial Intelligence. Stanford University, 2016. Acesso em: 28 set. 2024. Disponível em: https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0831fnl.pdf.

SURDEN, H. Artificial Intelligence and Law: An Overview. **Georgia State University Law Review**, v. 35, n. 4, p. 1305-1337, 2019. Acesso em: 24 mai. 2024. Disponível em: https://scholar.law.colorado.edu/faculty-articles/1234?utm_source=scholar.law.colorado.edu%2Ffaculty-articles%2F1234&utm_medium=PDF&utm_campaign=PDFCoverPages

TOLEDO, A. T.; MENDONÇA, M. A aplicação da inteligência artificial na busca de eficiência pela administração pública. **Revista do Serviço Público**, v. 74, n. 2, p. 410-438, 2023. Disponível em: https://revista.enap.gov.br/index.php/RSP/article/view/6829. Acesso em: 10 mar. 2025.

TONMOY, S. M.; FAISAL, F.; RAHMAN, T.; WANG, H.; ANANIADOU, S.; ALI, J.; FIROOZI, R.; MOHIUDDIN, T.; WAHAB, M.; FAZLI VAFA, P. **A comprehensive survey of hallucination mitigation techniques in large language models**. 2024. Disponível em: https://arxiv.org/abs/2401.01313. Acesso em: 01 out. 2024.

TORRES, R. C. L. **Leis de Licitações públicas comentadas**. 12. ed. rev., ampl. e atual. São Paulo: Ed. Juspodiym, 2021.

TRENTO, C. S. A. **Direito à explicação e decisões algorítmicas**: perspectivas a partir da Resolução n. 332/2020 do Conselho Nacional de Justiça. 2021. Disponível em: https://repositorio.ufsc.br/handle/123456789/219478. Acesso em: 03 out. 2024.

TRIBUNAL DE JUSTIÇA DE PERNAMBUCO. **Plano de Transformação Digital 2021–2026**: por um judiciário mais digital, transparente e humanizado. Recife: TJPE, 2022. Disponível em: https://portal.tjpe.jus.br/web/transparencia/plano-de-transforma%C3%A7%C3%A3o-digital. Acesso em: 26 mai. 2025.

_____. **Portal da Transparência do TJPE**. Recife: TJPE, 2025. Disponível em: https://portal.tjpe.jus.br/web/licitacoes-e-compras/sistema-licitacoes-atas. Acesso em 02 mar. 2024.

TURING, A. M. Computing Machinery and Intelligence. **Mind**, v. 59, n. 236, p. 433-460, 1950.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is All You Need. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30, 2017. **Proceedings** [...]. 2017. p. 5998-6008. Disponível em: https://www.codetds.com/article/15517262. Acesso em: 28 set. 2023.

XU, Z.; JAIN, S.; KANKANHALLI, M. **Hallucination is Inevitable**: An Innate Limitation of Large Language Models. 2024. Disponível em: https://arxiv.org/abs/2401.11817. Acesso em: 28 set. 2024.

ZDNET. **ChatGPT vs Microsoft Copilot vs Gemini: Which is the best AI chatbot?** 2024. Disponível em: https://www.zdnet.com. Acesso em: 04 jun. 2025.

3. INTELIGÊNCIA ARTIFICIAL GENERATIVA NAS LICITAÇÕES PÚBLICAS: REQUISITOS PARA ESTRUTURAÇÃO DOS *PROMPTS* DE DISPENSA DE LICITAÇÃO

3.1. INTRODUÇÃO

A crescente demanda por modernização e eficiência nos processos administrativos do setor público tem direcionado a atenção para tecnologias emergentes capazes de otimizar procedimentos burocráticos sem comprometer o rigor técnico e a conformidade legal exigidos pela legislação brasileira. Nesse contexto, a Inteligência Artificial (IA) e, mais especificamente a Inteligência Artificial Generativa (IAG) apresentam-se como ferramentas promissoras para transformar os diversos procedimentos administrativos, como os licitatórios, área crítica da gestão governamental que demanda rigor técnico, conformidade legal e transparência (Contini, 2024; Rossetti, Garcia, 2023).

Os processos licitatórios, regulamentados pela Lei nº 14.133/2021, representam um elemento fundamental da administração pública brasileira, constituindo o mecanismo oficial para aquisição de bens e serviços pelo Estado. Caracterizados por sua complexidade procedimental e pela necessidade de estrita observância aos princípios constitucionais da legalidade, impessoalidade, moralidade, publicidade e eficiência, esses processos frequentemente demandam considerável esforço administrativo e conhecimento técnico-jurídico especializado (Torres, 2021).

A implementação da IAG, particularmente por meio dos Grandes Modelos de Linguagem (LLMs), oferece oportunidades significativas para automatizar a elaboração de documentos, inclusive os licitatórios, potencialmente reduzindo erros, padronizando procedimentos e acelerando trâmites administrativos. No entanto, o sucesso dessa implementação depende fundamentalmente da qualidade das instruções fornecidas a esses sistemas, conhecidas como "*prompts* " (Liu et al*.,* 2023; Sahoo *et al.,* 2024). A engenharia de *prompts* – processo de desenvolvimento e refinamento de instruções para modelos de IAG – emerge, portanto, como um campo crucial para garantir que os sistemas generativos produzam documentos precisos, conformes e juridicamente fundamentados (Schulhoff et al., 2024).

O desafio central reside em traduzir os requisitos complexos e multidimensionais dos processos licitatórios em especificações técnicas que orientem adequadamente os modelos de

IAG. Esta tradução exige compreensão profunda tanto das capacidades e limitações tecnológicas dos LLMs quanto das especificidades jurídico-administrativas que caracterizam as contratações públicas no Brasil.

Neste contexto, o Termo de Referência emerge como artefato central para esta investigação. Definido pelo artigo 6º, inciso XXIII da Lei nº 14.133/2021 como documento técnico essencial que deve conter "parâmetros e elementos descritivos específicos" (Brasil, 2021), este instrumento apresenta características que o tornam ideal para explorar a aplicabilidade de IAG nos processos licitatórios. Sua natureza híbrida, combinando elementos técnicos descritivos com fundamentação jurídica, oferece ambiente de teste robusto para validar múltiplas dimensões da capacidade generativa dos modelos de linguagem, especialmente considerando os dez elementos obrigatórios estabelecidos pela legislação vigente.

Diante da convergência entre as potencialidades tecnológicas da IAG e as demandas específicas dos processos licitatórios, emerge a seguinte questão de pesquisa: **quais são os requisitos técnicos e normativos necessários para estruturar *prompts* otimizados que viabilizem a geração ética e transparente de artefatos de licitação na modalidade Dispensa de Licitação?** Esta interrogação articula-se em torno de três dimensões críticas: a primeira, relacionada às especificações técnicas que devem orientar a engenharia de *prompts* para garantir precisão e conformidade legal; a segunda, concernente aos marcos normativos e jurídicos que estabelecem os parâmetros de validade dos documentos licitatórios; e a terceira, referente aos aspectos éticos e de transparência que asseguram a implementação responsável da tecnologia no contexto da administração pública brasileira. Assim, a investigação desenvolvida no formato de ensaio teórico, tem como objetivo geral estabelecer as bases conceituais e técnicas necessárias para o desenvolvimento posterior de *prompts* customizados que atendam simultaneamente aos critérios de eficiência operacional e conformidade normativa.

A relevância desta investigação amplifica-se diante dos desafios identificados na literatura especializada sobre aplicação de IA no judiciário brasileiro. Souza *et al*. (2024) evidenciam que a falta de transparência e explicabilidade representa a principal preocupação dos pesquisadores, enquanto o fenômeno das alucinações – geração de conteúdo aparentemente coerente mas factualmente incorreto – constitui risco crítico para a confiabilidade dos sistemas automatizados (Tonmoy et al., 2024).

A escolha do Termo de Referência como objeto central desta pesquisa justifica-se por duas razões principais: sua importância técnica nos processos de Dispensa de Licitação e sua estrutura padronizada, que facilita o desenvolvimento de *prompts* específicos. Esta delimitação

estratégica permite validação rigorosa das técnicas propostas, servindo como modelo para futuras aplicações de IAG em outros documentos licitatórios.

A estrutura analítica do ensaio organiza-se em três seções complementares. A seção 3.2 examina em profundidade a IAG, explorando a arquitetura *transformer* que fundamenta os LLMs contemporâneos, as técnicas avançadas de engenharia de *prompts* aplicáveis ao contexto administrativo, e os desafios específicos relacionados ao controle de alucinações. A seção 3.3 aborda os fundamentos jurídico-administrativos das licitações públicas, estabelecendo o contexto normativo para aplicação de IAG e detalhando as especificidades da modalidade Dispensa de Licitação, com ênfase particular na caracterização do Termo de Referência como documento central da pesquisa. A seção 3.4 sintetiza os achados em um conjunto estruturado de requisitos para desenvolvimento de *prompts* eficazes.

O enfoque metodológico adotado combina revisão da literatura técnica sobre engenharia de *prompts* com análise jurídico-administrativa dos marcos normativos brasileiros, buscando identificar pontos de convergência entre as possibilidades tecnológicas e os requisitos específicos do domínio licitatório, tornando a abordagem interdisciplinar.

A contribuição esperada desta investigação articula-se em duas dimensões principais: metodológica, através da sistematização de critérios específicos para engenharia de *prompts* em contextos jurídico-administrativos; e prática, mediante o estabelecimento de diretrizes técnicas que orientem o desenvolvimento de soluções de IAG adequadas às especificidades das contratações públicas brasileiras.

3.2. INTELIGÊNCIA ARTIFICIAL GENERATIVA

A IAG tem emergido como um ramo revolucionário da inteligência artificial, destacando-se por sua capacidade de gerar novos conteúdos, como imagens, textos e música, simulando aspectos da criatividade humana (Takale; Mahalle; Sule, 2024). Essa tecnologia se diferencia de sistemas de IA tradicionais, que focam em tarefas preditivas e de classificação, enquanto a IAG busca replicar o processo criativo humano, gerando dados realistas semelhantes aos conjuntos de treinamento (Grossman; Grimm; Brown, 2023). Segundo Hessel e Lemes (2023), a IAG é capaz de simular a criatividade humana, gerando soluções originais com base em dados pré-existentes, o que a torna uma ferramenta poderosa em diversas áreas, incluindo as artes, *design*, e, mais recentemente, no contexto jurídico.

O conceito de modelos generativos remonta aos primórdios da pesquisa em IA, com foco inicial no desenvolvimento de algoritmos capazes de produzir novos pontos de dados similares aos encontrados em determinados conjuntos. Um dos primeiros exemplos foram as

cadeias de Markov, propostas pelo matemático russo Andrey Markov no final do século XIX, que descrevem sequências de eventos em que a probabilidade de cada ocorrência depende apenas do estado anterior (Takale; Mahalle; Sule, 2024).

Ao longo do século XX, a área de IA avançou com o desenvolvimento de sistemas especialistas baseados em regras e com a IA simbólica, pois, embora eficazes em domínios específicos, careciam da capacidade de gerar novo conteúdo ou aprender a partir de dados (Kissinger; Schmidt; Hottenlocher, 2021).

Além disso, uma mudança significativa ocorreu com a transição de abordagens convencionais para métodos baseados em aprendizado profundo na modelagem generativa. As redes neurais profundas, com múltiplas camadas de nós interconectados, permitiram modelar relações complexas e produzir saídas altamente realistas (Takale; Mahalle; Sule, 2024). O desenvolvimento da IAG foi impulsionado por diversas técnicas, incluindo as Redes Neurais Artificiais (RNAs) e os LLMs (Rossetti, Garcia, 2023). Entretanto, a técnica revolucionária foi a dos *transformers,* discutida adiante. O Quadro 6 lista as principais diferenças entre os modelos de IA e IAG.

Quadro 6: Diferenças entre modelos de inteligência artificial descritivos e generativos

	Treinamento	Modelo	Interface	Aplicações
Modelos Descritivos	Ocorre em bases de dados ligados a determinado domínio e voltados para propósitos específicos	Metodologias diversas baseadas em correlações estatísticas observadas nas bases de dados do domínio	Usuário deve adquirir conhecimento específico do sistema para realizar os comandos ou as requisições	Classificadores de dados que detectam padrões e fazem predições ou adotam decisões e fazem recomendações baseadas nas predições
Modelos Generativos	Ocorre em bases de dados enormes e generalizadas (e.g. conteúdo disponível na Internet); pode ser complementado com treinamento para aplicações específicas (*fine tunning*)	Metodologias e arquiteturas diversas, não determinísticas e recorrentes, de geração e seleção do conteúdo gerado otimizando aquele que mais se aproxima do conteúdo gerado por humanos	Usam grandes modelos de linguagem e processamento de linguagem natural, de modo que o usuário pode interagir e fazer requisições em linguagem ordinária e os outputs são de fácil compreensão	Geram conteúdo novo de texto, imagem, áudio, vídeo e códigos de programação indistinguíveis do conteúdo produzido por humanos

Fonte: CNJ (2024e, p. 26)

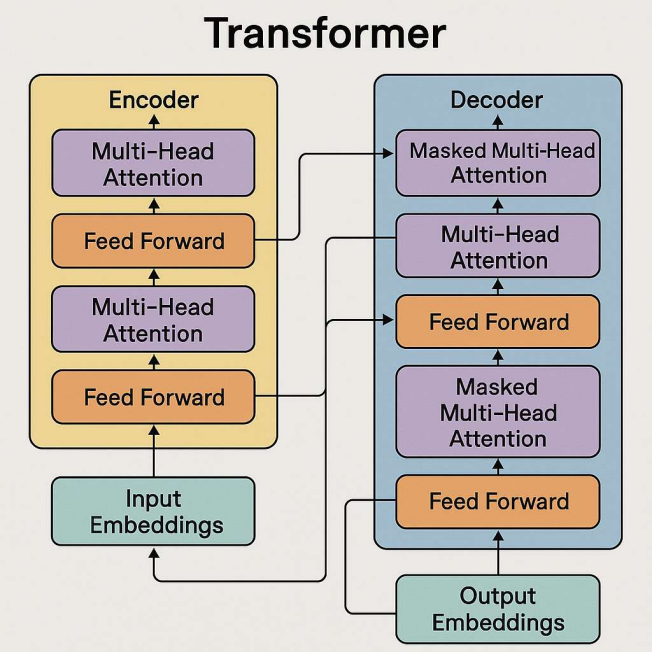
3.2.1. Transformer

O *transformer* é uma arquitetura de rede neural que desempenha um papel central no desenvolvimento de LLMs e, consequentemente, na IAG. Introduzido inicialmente por Vaswani *et al*. (2017), o *transformer* revolucionou o campo do processamento de linguagem natural (PLN) ao substituir arquiteturas tradicionais baseadas em redes neurais recorrentes (RNNs) e *long short-term memory* (LSTMs), que eram amplamente utilizadas até então. A

principal inovação do *transformer* está em seu mecanismo de *self-attention*, que permite ao modelo avaliar a importância de diferentes partes da entrada (seja uma sequência de texto ou outro tipo de dado) em paralelo, ao contrário das arquiteturas anteriores, que processavam os dados de maneira sequencial (Shen *et al.,* 2023).

A compreensão da arquitetura *transformer* torna-se fundamental para a aplicação eficaz de IAG em contextos administrativos, uma vez que suas características estruturais determinam diretamente as possibilidades e limitações dos modelos de linguagem utilizados. A figura 19 ilustra esquematicamente os componentes essenciais do mecanismo *transformer*, evidenciando como o processamento paralelo de informações e o sistema de atenção distribuída viabilizam a geração de texto contextualmente relevante e estruturalmente coerente - características indispensáveis para a elaboração de documentos licitatórios que devem atender simultaneamente a critérios de precisão técnica e conformidade normativa.

Figura 19: Mecanismo do *Transformer*

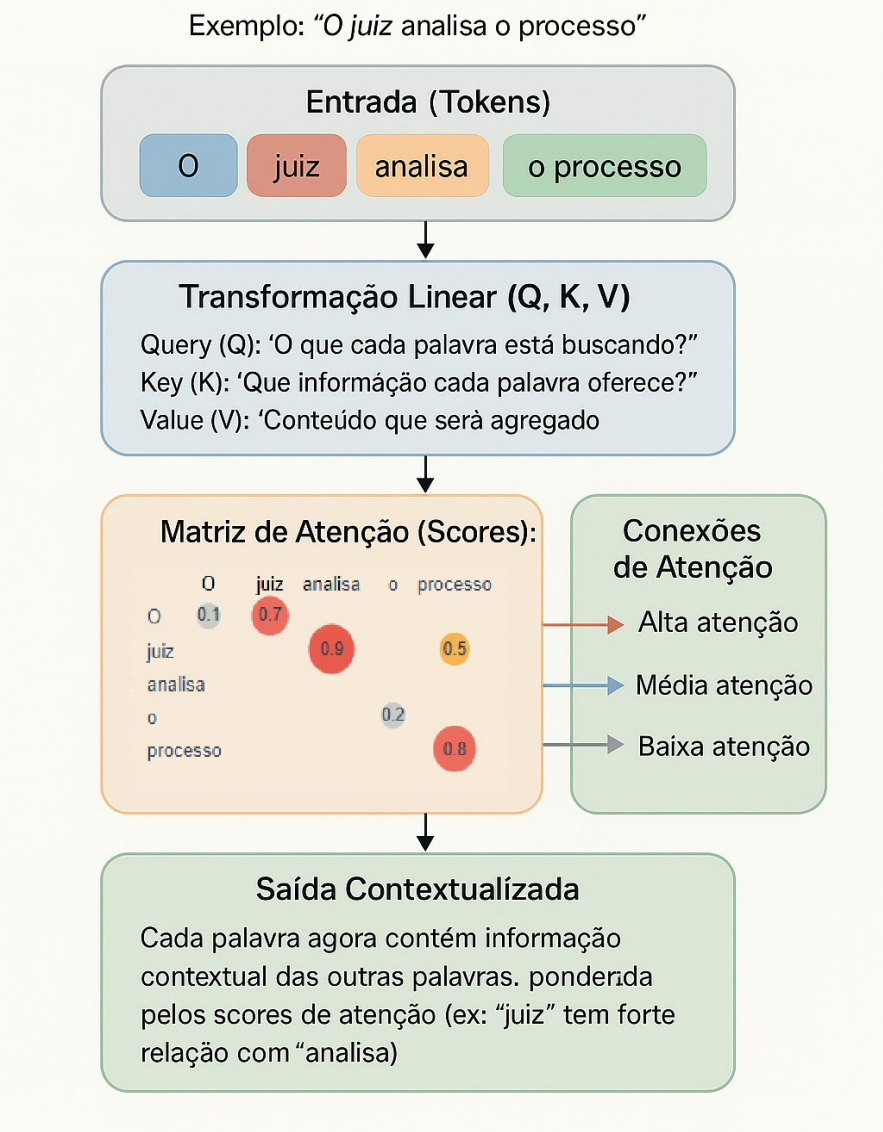


Fonte: Elaborado pela autora (2025)

O mecanismo de *self-attention* permite que o *transformer* capture de forma eficiente as dependências de longo alcance em um texto, reconhecendo a relação entre palavras ou frases que estão distantes na sequência, evidenciado visualmente na figura 20. Isso é crucial em tarefas de PLN, onde o contexto de palavras ou frases anteriores pode influenciar significativamente o significado de uma sentença subsequente. Além disso, o *transformer* é altamente escalável e permite o treinamento em grandes conjuntos de dados, o que possibilitou a criação de modelos

como o GPT-3, que possui 175 bilhões de parâmetros e é baseado nessa arquitetura, gerando novas possibilidades como a criação de imagens e a síntese de áudio, devido à sua capacidade de aprender padrões complexos em grandes volumes de dados (Jovanović *et al.,* 2022; Shen *et al.,* 2023).

Figura 20: Diagrama do mecanismo *Self-attention*

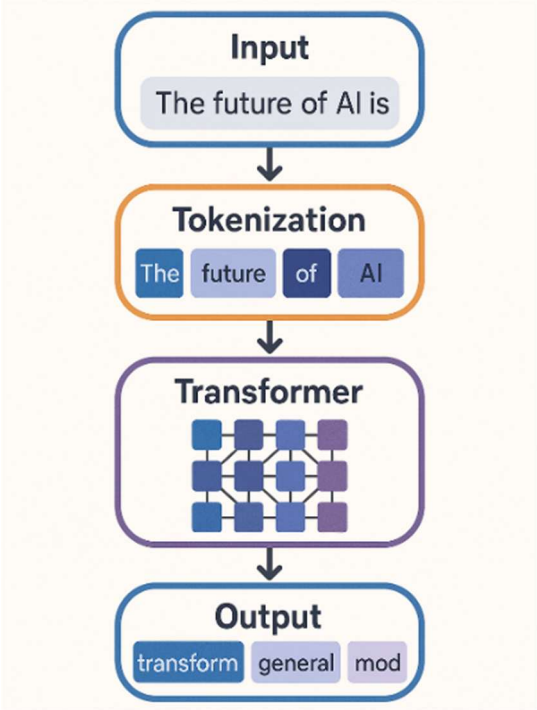


Fonte: Elaborado pela autora (2025)

Além do LLM, o *transformer* também foi adaptado para outras áreas da IAG, como a criação de imagens e a síntese de áudio. Como resultado, ele se tornou a base para diversos avanços na IA, incluindo o desenvolvimento de modelos como o GPT e o BERT, que são amplamente utilizados tanto em pesquisa quanto em aplicações práticas, como o ChatGPT (Shen *et al.,* 2023).

A implementação prática dos *transformers* em Grandes Modelos de Linguagem requer uma arquitetura específica que integre múltiplas camadas de processamento e mecanismos de atenção em uma estrutura coesa e escalável. A figura 21 apresenta a arquitetura típica de um LLM, demonstrando como os componentes fundamentais - desde a camada de *embedding* até os mecanismos de decodificação - se organizam para viabilizar o processamento de sequências textuais complexas. Esta estrutura arquitetônica é particularmente relevante para compreender como os modelos conseguem manter coerência contextual em documentos extensos, como os termos de referência e minutas contratuais característicos dos processos de Dispensa de Licitação, onde a preservação da consistência normativa e a precisão terminológica são requisitos indispensáveis.

Figura 21: Arquitetura do LLM.



Fonte: Elaborado pela autora (2025)

3.2.2. Prompts

Os *prompts* emergiram como uma técnica fundamental para estender as capacidades dos LLM e dos modelos de visão-linguagem (VLMs), permitindo sua adaptação para diversas tarefas sem modificar os parâmetros centrais do modelo (Liu et al., 2023). Ao contrário das abordagens tradicionais de aprendizado supervisionado, ou mesmo do ajuste fino de modelos, a engenharia de *prompts* permite que esses sistemas abordem tarefas específicas usando instruções cuidadosamente projetadas.

Diversos aspectos relevantes sobre a engenharia de *prompts* emergem, delineando um panorama atual e promissor dessa metodologia (Sahoo et al*.,* 2024). No contexto da administração pública, a aplicação eficaz de *prompts* torna-se essencial para a automação de processos burocráticos, garantindo precisão, transparência e conformidade legal na geração de documentos.

3.2.2.1. Evolução e Classificação dos *Prompts*

A evolução dos paradigmas de aprendizado em LLM representa um marco significativo na computação moderna. Inicialmente, os modelos eram estritamente supervisionados, exigindo grandes volumes de dados rotulados para treinar as redes neurais. No entanto, à medida que a quantidade de dados rotulados se mostrou insuficiente para diversas aplicações, especialmente em novos contextos, houve uma mudança para o paradigma de pré-treino e ajuste fino (Liu et al*.,* 2023).

A engenharia de *prompts* surge nesse contexto como uma evolução natural, onde a tarefa é reformulada para se alinhar com o treinamento original do modelo, eliminando a necessidade de reestruturar o modelo para cada nova função. Esta transição paradigmática fundamenta a taxonomia contemporânea dos *prompts* , que pode ser organizada conforme apresentado no quadro 7.

Quadro 7: Taxonomia contemporânea dos *prompts*

Dimensão	Classificação	Características	Aplicabilidade
Estrutura	Estáticos	*Templates* fixos, implementação simples	Documentos padronizados
	Dinâmicos	*Templates* personalizados por entrada	Processos complexos variáveis
Representação	Discretos	*Strings* de texto natural, interpretáveis	Interface humano-máquina
	Contínuos	Operação no espaço de *embeddings*	Otimização automática
Metodologia	*Zero-shot*	Sem exemplos de demonstração	Tarefas diretas e bem definidas
	Few-shot	Com poucos exemplos orientadores	Tarefas específicas do domínio
	Chain-of-Thought	Raciocínio passo-a-passo explícito	Problemas complexos multi-etapas

Fonte: Baseado em Liu *et al*. (2023) e Schulhoff *et al*. (2024)

A distinção entre *prompts* estáticos e dinâmicos revela-se fundamental para o contexto licitatório. Enquanto os estáticos utilizam *templates* fixos, sendo mais fáceis de implementar, mas menos adaptáveis, os dinâmicos geram *templates* personalizados para cada entrada, melhorando significativamente a contextualização (Liu et al., 2023). Esta característica torna-se especialmente relevante quando se considera a diversidade de objetos e modalidades contratuais presentes nas dispensas de licitação.

Paralelamente, a diferenciação entre *prompts* discretos e contínuos estabelece outro eixo crucial de classificação. Os discretos, representados por *strings* de texto natural, são mais facilmente interpretáveis pelos humanos, enquanto os contínuos operam diretamente no espaço de *embeddings*, oferecendo maior flexibilidade, mas sendo menos interpretáveis (Liu et al., 2023).

3.2.2.2. Técnicas Avançadas de Engenharia de *Prompts*

O desenvolvimento de técnicas avançadas de engenharia de *prompts* tem sido impulsionado pela necessidade de maior precisão, consistência e confiabilidade na interação com LLMs. Schulhoff *et al.* (2024) apresentam uma taxonomia abrangente dessas técnicas, organizando-as em categorias funcionais que atendem a diferentes necessidades operacionais. A figura 22 ilustra essa organização taxonômica.

Figura 22: Taxonomia de técnicas de engenharia de *prompts*



Fonte: Elaborado pela autora baseado em Schulhoff *et al.* (2024)

Chain-of-Thought (CoT) e Variações

A técnica *Chain-of-Thought* representa um avanço significativo na capacidade dos LLMs de abordar problemas complexos que requerem raciocínio multi-etapas. Conforme demonstrado por Wei *et al.* (2022), esta abordagem induz o modelo a articular explicitamente cada estágio intermediário do processo cognitivo, estabelecendo uma trilha de raciocínio que conecta premissas e conclusões de forma transparente.

Evoluções da técnica CoT incluem o *Tree of Thoughts* (ToT), que permite exploração de múltiplas ramificações de raciocínio simultaneamente, e o *Graph of Thoughts* (GoT), que modela interdependências complexas entre diferentes linhas de argumentação (Schulhoff et al*.,* 2024). Essas variações mostram-se particularmente úteis em cenários onde a análise de conformidade legal exige consideração simultânea de múltiplos requisitos normativos.

Técnicas de Ação e Iteração

As técnicas de ação e iteração representam uma categoria emergente que combina capacidades de raciocínio com mecanismos de refinamento contínuo. A metodologia *ReAct* (*Reasoning and Acting*) exemplifica essa abordagem ao integrar processos de raciocínio com ações específicas, permitindo que o modelo refine suas próprias respostas iterativamente (Yao et al*.,* 2022, conforme citado em Schulhoff et al*.,* 2024).

Complementarmente, a técnica *Self-Refine* implementa um mecanismo de autocorreção onde a IA revisa automaticamente suas respostas antes da apresentação final, garantindo maior conformidade legal e padronização dos documentos gerados (Madaan et al*.,* 2023, conforme citado em Schulhoff et al*.,* 2024).

Retrieval-Augmented Generation (RAG)

A técnica RAG emerge como solução fundamental para o desafio da atualização de conhecimento em LLMs. Esta abordagem permite que o modelo consulte bases de dados externas antes de gerar suas respostas, assegurando que os documentos licitatórios sejam embasados em fontes confiáveis e atualizadas (Lewis et al*.,* 2020).

O Quadro 8 apresenta uma síntese comparativa das principais técnicas de engenharia de *prompts* e suas aplicabilidades específicas no contexto licitatório.

Quadro 8: Técnicas de engenharia de *prompt*

Técnica	Princípio Operacional	Vantagens	Aplicação em Licitações
Chain-of-Thought	Raciocínio explícito passo-a-passo	Transparência, auditabilidade	Justificativas de dispensa, fundamentação legal
Few-shot Learning	Aprendizado por exemplos	Padronização, consistência	*Templates* de documentos, minutas contratuais
ReAct	Raciocínio + ação iterativa	Autocorreção, precisão	Verificação de conformidade, validação documental
Self-Refine	Autorrevisão sistemática	Qualidade, confiabilidade	Controle de qualidade, revisão automática
RAG	Consulta às bases externas	Atualização, precisão factual	Verificação normativa, consulta jurisprudencial
Multi-Agent	Múltiplas perspectivas	Robustez, completude	Análise multidisciplinar, pareceres técnicos

Fonte: Baseado em Schulhoff *et al*. (2024) e literatura correlata

3.2.2.3. Parâmetros de Configuração de *Prompts*

Persona

Uma estratégia promissora na engenharia de *prompts* consiste na incorporação de *personas*, que podem ser definidas como perfis contextuais (reais ou fictícios) atribuídos ao modelo de linguagem ou ao usuário, com o objetivo de orientar a geração textual de forma mais alinhada ao domínio da tarefa. No contexto da IAG, as abordagens baseadas em persona têm se dividido em dois grandes eixos: *LLM Role-Playing* e *LLM Personalization* (Tseng et al., 2024).

No eixo *role-playing*, o modelo assume uma persona específica, com características e valores associados a um papel ou profissão; por exemplo, juiz, engenheiro, analista jurídico etc., o que influencia diretamente o estilo, o conteúdo e a estrutura das respostas. Essa técnica tem se mostrado eficaz em ambientes de simulação jurídica e diagnósticos médicos, pois permite ao modelo comportar-se com coerência e consistência em cenários específicos (Park et al., 2023; Tang et al., 2023).

Por outro lado, o eixo *personalization* baseia-se no perfil do usuário, como histórico de interações, preferências ou dados demográficos, para adaptar dinamicamente as respostas ao

contexto individual. Embora mais difundida em sistemas de recomendação e diálogos personalizados, essa técnica tem potencial de aplicação em ambientes administrativos, como a geração de artefatos licitatórios, ao considerar, por exemplo, o padrão textual ou normativo da instituição pública a que se destina o documento (Abbasian et al., 2024; Zhang et al., 2023). O quadro 9 faz um comparativo entre essas duas abordagens, com destaque para suas definições, aplicações, vantagens e riscos associados, além da aplicabilidade prática ao contexto das licitações públicas por dispensa.

Quadro 9: Comparativo entre as abordagens *Role-Playing e Personalization* na Engenharia de *Prompts* com Persona

Critério	LLM *Role-Playing*	LLM *Personalization*
Definição	O modelo assume uma persona (papel ou função) específica, com características pré-definidas.	O modelo adapta suas respostas com base no perfil e preferências do usuário.
Foco do Contexto	Externo ao usuário – centrado no papel que o modelo desempenha.	Interno ao usuário – centrado em suas necessidades, histórico e estilo.
Aplicações típicas	Simulações jurídicas, diagnósticos médicos, agentes em jogos, julgamentos automatizados.	Sistemas de recomendação, tutores educacionais, assistentes pessoais, geração de relatórios customizados.
Exemplo no contexto jurídico	"Você é um analista jurídico do Tribunal e deve elaborar uma minuta de dispensa com base no art. 75..."	"Considere que o usuário elabora frequentemente contratos de serviços contínuos para o Tribunal..."
Vantagens	Gera respostas mais consistentes com o papel atribuído; melhora a coerência estilística e argumentativa.	Ajuste fino às necessidades institucionais; maior aderência ao estilo textual desejado pelo órgão.
Riscos e desafios	Possibilidade de estereótipos ou enviesamentos da persona simulada; risco de comportamento destrutivo.	Riscos à privacidade e à ética na coleta e uso de dados pessoais / institucionais.
Aplicabilidade na pesquisa	Pode ser utilizada para simular o comportamento de diferentes atores do processo licitatório.	Permite gerar documentos adaptados ao perfil administrativo do Tribunal, aumentando a acurácia normativa.

Fonte: Baseado em Abbasian *et al*. (2024), Chen *et al*. (2023), Park *et al*. (2023) e Tseng *et al*. (2024)

Dessa forma, o uso de personas em *prompts* atua como uma camada de contexto que orienta o modelo a produzir saídas mais adequadas, tanto em termos de conteúdo quanto de forma. Em ambientes sensíveis como o da administração pública, a atribuição de uma persona pode contribuir para garantir maior conformidade legal e coesão técnico-normativa, além de mitigar desvios de estilo e conteúdo. Essa técnica é especialmente útil quando combinada com

outras estratégias, como *chain-of-thought* ou *retrieval-augmented generation*, uma vez que amplia a ancoragem contextual da resposta.

Além disso, estudos têm apontado que a adoção de personas pode influenciar comportamentos emergentes entre agentes LLM em ambientes colaborativos: em cenários multiagente, por exemplo, o uso de personas distintas permite simular interações complexas, com comportamentos cooperativos, conformistas ou até destrutivos – fenômenos que ilustram o potencial e os riscos associados à técnica (Chen et al., 2023; Guo et al., 2024). Assim, a engenharia de *prompts* baseada em persona não apenas melhora a qualidade da resposta gerada, mas também permite controlar aspectos éticos, alinhamento normativo e simulações interativas de cenários administrativos realista.

Temperatura: controle de aleatoriedade e criatividade na geração textual

No processo de geração de texto por LLMs, o parâmetro denominado temperatura (*temperature*) atua como um modulador da aleatoriedade na escolha de palavras. Trata-se de um parâmetro responsável por ajustar a distribuição de probabilidades das próximas palavras a serem geradas (Peeperkorn et al, 2024). Valores baixos (próximos de zero) reduzem a variabilidade do *output*, favorecendo respostas mais previsíveis e conservadoras; já valores mais altos aumentam a entropia, gerando saídas mais criativas, porém menos consistentes (IBM, 2024).

Embora frequentemente rotulado como "o parâmetro da criatividade", estudos recentes sugerem que a relação entre temperatura e criatividade é mais frágil do que se presume. Peeperkorn *et al*. (2024) realizaram uma investigação empírica com o modelo LLAMA 2-CHAT, evidenciando que, apesar de existir uma correlação positiva entre temperatura e novidade dos textos, também há uma correlação negativa com a coerência, indicando um *trade-off*[5] entre inovação e compreensibilidade. Ademais, não foram encontradas associações significativas com os critérios de coesão e tipicidade, o que enfraquece a noção de que a elevação da temperatura por si só confere caráter criativo aos *outputs*.

No contexto da engenharia de *prompts* para licitações públicas, o ajuste adequado da temperatura torna-se relevante para equilibrar precisão normativa e capacidade de adaptação textual. Temperaturas mais baixas são indicadas para redação de minutas, pareceres e termos

[5] *Trade-off* é um termo amplamente utilizado nas ciências exatas, computacionais e econômicas para descrever uma situação de compensação mútua, em que a melhoria de um atributo resulta na deterioração de outro. No contexto dos LLMs, refere-se ao equilíbrio entre originalidade e clareza: quanto maior a liberdade criativa (temperatura), menor tende a ser a coerência do texto gerado (IBM, 2024; Peeperkorn et al., 2024).

de referência, enquanto configurações medianas podem ser exploradas em processos iterativos de *brainstorm* de justificativas ou simulações de perfis institucionais. Configurações mais altas devem ser aplicadas com cautela, especialmente em tarefas que exigem aderência estrita a marcos legais e padronizações.

A seguir, o quadro 10 sintetiza a relação das diferentes faixas de temperatura com suas características, aplicações e exemplos práticos no domínio jurídico-administrativo.

Quadro 10*:* Efeitos da Temperatura em Modelos de Linguagem

Faixa de Temperatura	**Características do *Output***	**Casos de Uso**	**Exemplos no Contexto Jurídico**
Baixa (como 0.2)	Previsível, focado, consistente	Redação formal, respostas normativas, FAQs	Elaboração de minutas padrão, pareceres técnicos, termos de referência
Média (como 0.7)	Balanceado entre criatividade e coerência	Diálogos, explorações, simulações	Simulações de persona institucional, *brainstorm* de justificativas legais
Alta (como 1.0+)	Criativo, variado, menos previsível	Produção artística, narrativas, ideação	Sugestões de alternativas inovadoras em cláusulas contratuais

Fonte: Baseado em IBM (2024) e Peeperkorn *et al*. (2024)

A configuração do parâmetro temperatura constitui decisão técnica fundamental para aplicações de IAG em contextos regulamentados. Para o domínio das licitações públicas, a literatura especializada indica a necessidade de temperaturas baixas por três razões teóricas centrais: i) a natureza altamente normatizada dos processos licitatórios exige previsibilidade e consistência nas saídas geradas, características favorecidas por configurações conservadoras do parâmetro (Peeperkorn et al., 2024); ii) a redução da aleatoriedade contribui para a mitigação de alucinações, limitação crítica inerente aos LLMs (Xu et al., 2024); iii) a configuração de baixa entropia alinha-se aos requisitos de auditabilidade e transparência estabelecidos pela Resolução CNJ nº 615/2025.

3.2.2.4. Integração de Técnicas e Arquiteturas Híbridas

A fronteira atual da engenharia de *prompts* caracteriza-se pela integração sistemática de múltiplas técnicas em arquiteturas híbridas que maximizam as potencialidades individuais de cada abordagem. Nesse contexto, algumas técnicas de engenharia de *prompts* mostram-se particularmente adequadas para aplicações no domínio das licitações públicas, cada uma atendendo a requisitos específicos do contexto administrativo.

A **Chain-of-Thought (CoT)** emerge como técnica fundamental para estruturação do raciocínio jurídico-administrativo, pois como demonstram Wei *et al*. (2022), esta abordagem permite explicitar o processo de fundamentação, característica essencial para documentos que devem resistir ao escrutínio de órgãos de controle e auditoria.

A **Retrieval-Augmented Generation (RAG)** apresenta-se como solução técnica para o requisito de atualização normativa, pois com evidenciam Lewis *et al*. (2020), esta técnica permite integrar bases de conhecimento externas, mitigando limitações temporais dos modelos de linguagem e garantindo aderência ao arcabouço legal vigente.

O **Self-Refine** se constitui como mecanismo de controle de qualidade automatizado, pois como demonstram Madaan *et al*. (2023), esta técnica reduz inconsistências e aprimora a precisão textual através de processos iterativos de autocorreção, aspecto crítico para documentos administrativos de alto rigor técnico.

Já **Persona** representa parâmetro de configuração transversal que define o perfil contextual do modelo; Tseng *et al*. (2024) classificam esta abordagem como elemento de contextualização que permeia todas as demais técnicas, orientando estilo, terminologia e nível de detalhamento adequados ao domínio específico.

Por fim, a utilização de uma **Temperatura** baixa (0.2 ou inferior) constitui parâmetro essencial para o contexto licitatório, uma vez que como Peeperkorn *et al*. (2024) apontam, configurações conservadoras favorecem a coerência e previsibilidade em detrimento da criatividade, características alinhadas às exigências de conformidade normativa e auditabilidade do setor público. Esta configuração reduz significativamente a incidência de alucinações (Xu et al., 2024), aspecto crítico para documentos que fundamentam gastos públicos e decisões contratuais.

Assim, a combinação dessas técnicas atende aos requisitos multidimensionais do contexto licitatório: fundamentação legal (CoT), precisão factual (RAG), controle de qualidade (Self-Refine) e adequação contextual (Persona e Temperatura), estabelecendo as bases teóricas para implementação de soluções robustas e confiáveis.

3.2.3. Alucinações

A alucinação em modelos de IAG tornou-se um tópico de crescente importância à medida que esses modelos são cada vez mais aplicados em contextos críticos, como o setor jurídico, a saúde e outras áreas sensíveis. A alucinação refere-se à capacidade dos modelos de gerar conteúdo que parece coerente e factível, mas que, na realidade, é falso ou inventado. Este fenômeno desafia não apenas a precisão das respostas fornecidas pela IAG, mas também levanta

preocupações sobre o uso confiável desses sistemas em ambientes onde a veracidade da informação é crucial (Rawte et al*.,* 2023).

Historicamente, os LLMs evoluíram com base em vastos volumes de dados textuais coletados da Internet, muitos dos quais são imperfeitos, imprecisos ou desatualizados. Como resultado, esses modelos são treinados para identificar padrões nesses dados e gerar texto coerente, mas não são necessariamente programados para verificar a precisão factual das informações que produzem. A alucinação é, portanto, um subproduto inevitável dessa forma de treinamento, uma vez que os LLMs, como o GPT-3 e suas versões subsequentes, não têm uma compreensão "real" do mundo, mas apenas reproduzem padrões linguísticos (Shen et al*.,* 2023).

Uma das razões pelas quais a alucinação ocorre nos LLMs está relacionada à atenção diluída no mecanismo de *self-attention* do *transformer*, o núcleo da arquitetura dos LLMs. Conforme a sequência de texto se torna mais longa, o modelo tende a perder a capacidade de relacionar adequadamente todas as partes do *input*, o que pode resultar em respostas que desviam do contexto correto. Além disso, há um efeito conhecido como *exposure bias*, onde a discrepância entre os dados usados no treinamento supervisionado e o ambiente real de geração de texto pode induzir alucinações (Xu et al*.,* 2024).

A alucinação pode ser classificada em dois tipos principais: intrínseca e extrínseca. A alucinação intrínseca ocorre quando o modelo gera informações que são inconsistentes ou contraditórias com os dados de entrada fornecidos, como instruções ou *prompts* . Já a alucinação extrínseca surge quando o modelo produz respostas que não podem ser verificadas com base no conteúdo existente do *input*, ou seja, quando ele extrapola informações não factuais ou irreais (Tonmoy et al*.,* 2024). Essa distinção é fundamental para compreender como mitigar esses problemas em diferentes contextos.

Diferentes estratégias têm sido propostas para mitigar a alucinação. Uma delas é o uso de métodos de geração aumentada por recuperação (RAG), que combina os modelos generativos com bases de conhecimento externas, como bancos de dados jurídicos ou enciclopédias. Isso permite que o modelo verifique suas respostas contra fontes confiáveis antes de gerá-las, reduzindo a probabilidade de alucinação (Tonmoy et al*.,* 2024). Além disso, técnicas de refinamento e *feedback* também têm sido exploradas para permitir que os modelos corrijam seus próprios erros durante o processo de geração, proporcionando saídas mais precisas e confiáveis (Shen et al*.,* 2023).

Outro aspecto crucial da alucinação é a falta de atualizações em tempo real nos modelos generativos. Como a maioria dos LLMs é treinada *em corpora* que não são atualizados regularmente, os modelos podem fornecer informações desatualizadas ou incorretas em tópicos

que mudam rapidamente, como regulamentações legais ou descobertas científicas recentes. Isso é particularmente problemático em áreas como o direito e a medicina, onde a precisão e a atualidade das informações são fundamentais (Shen et al*.,* 2023).

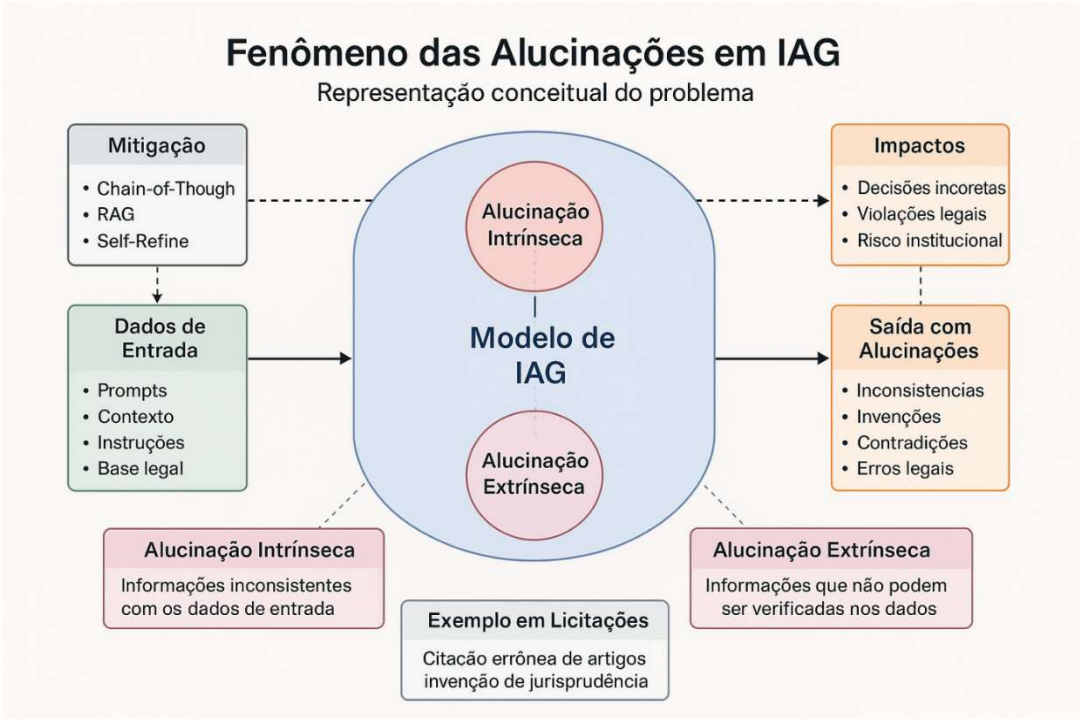
No contexto jurídico, a alucinação assume um risco ainda maior, dado o impacto potencial de informações erradas em decisões legais e na administração da justiça. Como observado por Contini (2024), o uso de IAG no setor legal tem sido promissor, mas as alucinações podem gerar sérios problemas, especialmente quando os modelos criam respostas falsas sobre leis ou casos inexistentes. Isso levanta a necessidade urgente de mecanismos rigorosos de validação para evitar decisões judiciais baseadas em fatos incorretos (Contini, 2024).

O fenômeno das alucinações em modelos de IAG pode ser melhor compreendido por meio de uma representação visual que demonstre como informações aparentemente coerentes podem divergir da realidade factual, comprometendo a confiabilidade dos sistemas automatizados. A figura 23 ilustra esquematicamente o processo pelo qual os LLMs podem gerar conteúdo factualmente incorreto, mas estruturalmente plausível, evidenciando os pontos críticos onde a supervisão humana se torna indispensável. Esta visualização é particularmente relevante para o contexto das licitações públicas, onde a geração de informações imprecisas sobre dispositivos legais, valores contratuais ou especificações técnicas pode invalidar completamente um processo administrativo, resultando em prejuízos tanto para a administração pública quanto para os potenciais fornecedores.

Em termos de governança, a alucinação também levanta questões éticas. O uso de IA em áreas críticas exige altos padrões de responsabilidade, mas como se pode responsabilizar um sistema que não é capaz de "entender" o que está gerando? Esta questão está no centro das discussões sobre a regulamentação da IA, e há consenso de que a supervisão humana é indispensável no uso dessas ferramentas, especialmente em ambientes onde erros podem ter consequências graves (Contini, 2024).

A transparência quanto ao uso de IAG também é uma questão debatida. Enquanto algumas jurisdições, como certos tribunais nos Estados Unidos, exigem a divulgação do uso de IAG em documentos legais, outras, como a abordagem inglesa, enfatizam a responsabilidade profissional do usuário em garantir a precisão do conteúdo, independentemente de sua origem (Grossman; Grimm; Brown, 2023).

Figura 23: Fenômeno das alucinações em IAG



Fonte: Elaborado pela autora (2025)

Apesar dessas inovações, as soluções propostas até agora são apenas parcialmente eficazes. Muitos dos métodos de mitigação ainda dependem fortemente da qualidade e abrangência das bases de dados externas, o que significa que, se essas fontes forem limitadas ou enviesadas, os problemas de alucinação podem persistir. Além disso, os modelos de IA têm uma capacidade limitada de raciocinar ou entender o mundo de maneira semelhante aos humanos, o que restringe sua habilidade de identificar e corrigir alucinações de forma autônoma (Rawte et al., 2023).

Em conclusão, a alucinação em modelos de IA generativa é um fenômeno inevitável, devido à natureza probabilística e aos métodos de treinamento desses sistemas. Embora as técnicas de mitigação estejam avançando, elas ainda estão longe de resolver completamente o problema. A aplicação responsável da IA no setor jurídico e em outros domínios críticos dependerá, em última análise, de uma combinação de ferramentas técnicas para reduzir a alucinação e de uma supervisão humana robusta para garantir que as respostas geradas sejam precisas e confiáveis.

Estas considerações são particularmente relevantes quando se deseja propor diretrizes para mitigação de riscos relacionados a alucinações e imprecisões na geração automatizada de artefatos de licitação.

3.3. FUNDAMENTOS DA LICITAÇÃO

3.3.1. Conceito, Natureza Jurídica e Fundamentos Constitucionais

A licitação pode ser compreendida como um procedimento administrativo que estabelece uma ligação entre a pretensão contratual da Administração e o fornecedor que irá atendê-la. Como define Torres (2021), é o trajeto procedimental que liga a pretensão contratual ao fornecedor escolhido para atendê-la, sendo o caminho determinado pelo legislador para a escolha do fornecedor apto a atender a necessidade administrativa.

Em termos práticos, assim como um particular que deseja adquirir um produto pesquisa preços, analisa custos e avalia sua disponibilidade financeira antes de selecionar a melhor proposta no mercado, a Administração Pública precisa realizar ações semelhantes. No entanto, em razão da indisponibilidade do interesse público e da necessidade de consagrar a isonomia e a impessoalidade, o legislador estabeleceu procedimentos formais prévios para realizar essa contratação (Torres, 2021).

O texto constitucional estabelece importantes diretrizes para as licitações públicas, notadamente através de três dispositivos principais (Brasil, 2024):

- Artigo 22, XXVII - Define a competência privativa da União para legislar sobre normas gerais de licitação;
- Artigo 37, XXI - Estabelece a obrigatoriedade da licitação pública;
- Artigo 173, § 1°, III - Dispõe sobre o estatuto jurídico das empresas estatais.

3.3.2. Lei nº 14.133/2021 - Nova Lei de Licitações

A nova Lei de Licitações e Contratos Administrativos (Lei 14.133/2021), promulgada em primeiro de abril de 2021, representa um importante marco de modernização legislativa no âmbito das contratações públicas brasileiras. Esta norma jurídica realizou a consolidação normativa de três instrumentos regulatórios anteriores: a legislação tradicional de licitações (Lei 8.666/1993), o normativo referente à modalidade pregão (Lei 10.520/2002) e o dispositivo legal que estabeleceu o Regime Diferenciado de Contratações Públicas (Lei 12.462/2011). A unificação destes três pilares legislativos em um único diploma legal busca promover maior racionalidade e eficiência no sistema de compras governamentais brasileiro, constituindo significativo avanço na sistematização do ordenamento jurídico administrativo nacional.

3.3.3. Modalidades de Compras Diretas

A Lei 14.133/2021 prevê hipóteses em que a Administração Pública pode realizar contratação direta, dispensando o processo licitatório formal. As duas modalidades principais são Dispensa de Licitação e inexigibilidade de licitação.

A Inexigibilidade de Licitação (Art. 74) ocorre quando não há possibilidade de competição, seja pela exclusividade do fornecedor ou pela natureza do serviço contratado. Exemplos incluem os serviços técnicos especializados de natureza intelectual com notória especialização (art. 74, inciso III); a contratação de fornecedor exclusivo (art. 74, inciso I) ou o credenciamento de fornecedores (art. 74, inciso IV).

A Dispensa de Licitação (Art. 75) ocorre quando a competição é possível, mas, por razões justificáveis, o legislador autoriza a não realização do certame. São exemplos a contratação de pequeno valor (art. 75, incisos I e II); situações emergenciais (art. 75, inciso VIII) ou aquisição de bens ou serviços em consórcio público (art. 75, inciso XX).

Mesmo sem o rito formal de uma licitação, a contratação direta exige justificativa, pesquisa de preços e parecer jurídico, garantindo transparência e economicidade. Estes procedimentos, ainda que simplificados em comparação com o processo licitatório completo, representam o contexto particular para o qual os *prompts* customizados serão desenvolvidos nesta pesquisa, conforme estabelecido no terceiro objetivo específico da dissertação.

3.3.4. Termo de Referência: Artefato Central da Pesquisa

O Termo de Referência (TR) constitui documento técnico essencial no âmbito das contratações públicas, definido pelo artigo 6º, inciso XXIII da Lei nº 14.133/2021 como o "documento necessário para a contratação de bens e serviços, que deve conter parâmetros e elementos descritivos específicos" (Brasil, 2021). Este dispositivo legal estabelece dez elementos obrigatórios que devem compor o documento: (a) definição do objeto; (b) fundamentação da contratação; (c) descrição da solução; (d) requisitos da contratação; (e) modelo de execução; (f) modelo de gestão do contrato; (g) critérios de medição e pagamento; (h) forma e critérios de seleção do fornecedor; (i) estimativas do valor da contratação; e (j) adequação orçamentária.

Para uma compreensão mais detalhada de como cada um desses elementos se articula e de que forma a IAG pode ser aplicada em sua estruturação, o Quadro 11 apresenta o detalhamento dos dez elementos obrigatórios previstos na Lei nº 14.133/2021, ressaltando sua relevância no contexto da engenharia de *prompts* para artefatos de licitação.

Quadro 11: Detalhamento dos dez elementos obrigatórios do Termo de Referência (Art. 6º, XXIII, Lei nº 14.133/2021)

Elemento Obrigatório	Descrição	Relevância para a IAG
(a) Definição do Objeto	Especifica o que será contratado.	A IAG pode auxiliar na redação clara e precisa do objeto, garantindo que seja abrangente e sem ambiguidades, com base em exemplos e normas.
(b) Fundamentação da Contratação	Justifica a necessidade da contratação.	A IAG pode estruturar a argumentação jurídica e administrativa, conectando a necessidade aos objetivos institucionais e legais, utilizando a técnica *Chain-of-Thought* para explicitar o raciocínio.
(c) Descrição da Solução	Detalha as características técnicas do bem ou serviço.	A IAG pode gerar descrições técnicas padronizadas e completas, minimizando erros e omissões, a partir de dados de entrada e referências existentes.
(d) Requisitos da Contratação	Define as condições que o fornecedor e o objeto devem atender.	A IAG pode ajudar a listar e organizar os requisitos de forma sistemática e conforme a legislação, garantindo a conformidade normativa.
(e) Modelo de Execução	Descreve como o contrato será executado.	A IAG pode auxiliar na elaboração de fluxos e procedimentos para a prestação do serviço ou entrega do bem, assegurando consistência e padronização.
(f) Modelo de Gestão do Contrato	Define como o contrato será acompanhado e fiscalizado.	A IAG pode gerar seções sobre responsabilidades, indicadores e mecanismos de controle, contribuindo para a transparência e a auditabilidade.
(g) Critérios de Medição e Pagamento	Estabelece como os serviços ou bens serão medidos e pagos.	A IAG pode formular cláusulas financeiras claras e em conformidade com as normas contábeis e fiscais, com base em modelos predefinidos.
(h) Forma e Critérios de Seleção do Fornecedor	Indica a modalidade de contratação e os critérios de julgamento.	A IAG pode auxiliar na descrição dos requisitos de habilitação e nas regras de seleção, promovendo a isonomia e a impessoalidade.
(i) Estimativas do Valor da Contratação	Apresenta o valor de referência para a contratação.	A IAG pode ajudar a consolidar dados de pesquisa de preços e justificar a estimativa de custo, utilizando a técnica *Retrieval-Augmented Generation* (RAG) para consultar bases de dados externas.
(j) Adequação Orçamentária	Comprova a disponibilidade de recursos.	A IAG pode auxiliar na redação da declaração de adequação orçamentária, garantindo a conformidade legal e a responsabilidade fiscal.

Fonte: Baseado no Art. 6º, XXIII, da Lei nº 14.133/2021 e na literatura sobre IAG.

A complexidade estrutural e a densidade técnico-jurídica do TR o posicionam como artefato ideal para investigação da aplicabilidade de IAG nos processos licitatórios por quatro razões fundamentais. Primeiro, sua natureza híbrida combina elementos técnicos descritivos com fundamentação jurídica, exigindo tanto precisão normativa quanto clareza expositiva,

características que permitem avaliar múltiplas dimensões da capacidade generativa dos modelos de linguagem.

Segundo, a padronização relativa de sua estrutura, estabelecida pela legislação, oferece *framework* estruturado que facilita a implementação de técnicas como *CoT*, permitindo que o modelo siga sequências lógicas previsíveis na construção do documento. Simultaneamente, a necessidade de adaptação às especificidades de cada contratação demanda flexibilidade contextual, testando a capacidade de personalização dos *prompts* desenvolvidos.

Terceiro, o TR representa documento de alto impacto no processo licitatório, uma vez que sua qualidade determina diretamente a eficácia da contratação subsequente. Torres (2021) destaca que deficiências na elaboração do TR constituem uma das principais causas de insucesso em processos licitatórios, seja por especificações inadequadas, seja por fundamentação jurídica insuficiente. Esta criticidade torna o documento um caso de teste rigoroso para avaliar a confiabilidade da automação baseada em IAG.

Quarto, sua utilização na modalidade Dispensa de Licitação apresenta volume significativo no setor público brasileiro, representando contexto empírico adequado para validação experimental. A representatividade quantitativa desta modalidade, combinada com a complexidade técnica do documento, proporciona ambiente controlado e relevante para demonstrar a aplicabilidade prática da pesquisa.

Assim, a escolha do TR como foco único desta investigação permite aprofundamento analítico que seria inviável em escopo mais amplo. Esta delimitação estratégica possibilita validação rigorosa das técnicas de engenharia de *prompts* propostas, estabelecendo precedente metodológico para futuras expansões da aplicação de IAG em outros artefatos licitatórios.

3.4. CONSIDERAÇÕES FINAIS

A investigação dos requisitos necessários para estruturação de *prompts* otimizados para geração de artefatos licitatórios revelou um conjunto complexo de especificações técnicas, jurídicas e éticas que devem ser rigorosamente observadas para viabilizar a implementação responsável de IAG no contexto das contratações públicas. A análise desenvolvida demonstra que o sucesso dessa implementação depende fundamentalmente da convergência entre precisão técnica, conformidade normativa e transparência operacional.

A arquitetura *transformer* estabeleceu as bases tecnológicas para o processamento de linguagem natural em larga escala, viabilizando modelos capazes de compreender e gerar conteúdo contextualmente relevante. Contudo, as limitações intrínsecas desses sistemas,

particularmente o fenômeno das alucinações identificado por Xu *et al*. (2024), exigem protocolos rigorosos de mitigação de riscos e supervisão humana especializada.

As técnicas de engenharia de *prompts* analisadas – *CoT, RAG, Self-Refine* e configuração de Persona – oferecem instrumentos complementares para estruturar soluções robustas e confiáveis. A configuração de temperatura baixa emerge como parâmetro crítico para garantir a previsibilidade e conformidade exigidas pelo contexto administrativo, conforme evidenciado por Peeperkorn *et al*. (2024).

A seleção do TR como artefato central desta pesquisa fundamenta-se em sua complexidade técnico-jurídica e representatividade no contexto das dispensas de licitação na administração pública. A estrutura híbrida deste documento, combinando elementos descritivos e fundamentação legal, proporciona ambiente de teste adequado para validar a eficácia das técnicas identificadas, especialmente considerando os dez elementos obrigatórios estabelecidos pelo artigo 6º, inciso XXIII da Lei nº 14.133/2021.

Os requisitos identificados para estruturação de *prompts* eficazes contemplam cinco dimensões essenciais: (i) fundamentação jurídica sólida, baseada na legislação vigente e na jurisprudência administrativa; (ii) estrutura lógica sequencial, que explicite o encadeamento do raciocínio através de técnicas como CoT; (iii) mecanismos de verificação e atualização de informações, operacionalizados através de RAG; (iv) protocolos de controle de qualidade, implementados via *Self-Refine*; e (v) contextualização adequada ao domínio específico, através de configuração apropriada de Persona.

Complementarmente, a supervisão humana emerge como elemento indispensável nesse processo. A complexidade e as implicações jurídicas dos documentos licitatórios exigem que profissionais especializados mantenham papel central na validação e revisão dos artefatos gerados por IAG, estabelecendo um modelo de colaboração humano-máquina que potencialize as capacidades de ambos os agentes, sem abdicar da responsabilidade e do julgamento profissional.

Destaca-se que o desenvolvimento e a implementação de IAG no contexto licitatório devem ser orientados não apenas por considerações de eficiência operacional, mas também por princípios éticos fundamentais, como transparência, equidade e responsabilidade. A otimização dos processos administrativos por meio dessas tecnologias deve servir ao propósito maior de fortalecer a integridade das contratações públicas, promovendo o uso responsável e eficiente dos recursos públicos em benefício da sociedade.

Cabe ressaltar que, embora promissora, a aplicação de IAG nos processos licitatórios ainda se encontra em estágio incipiente, demandando investigações empíricas adicionais para

validar a eficácia dos requisitos e técnicas identificados neste estudo teórico. Futuras pesquisas poderiam beneficiar-se de abordagens experimentais, que implementem e avaliem os *prompts* estruturados em ambientes controlados, permitindo mensurar seu desempenho em termos de precisão, conformidade legal e eficiência processual.

Por fim, reconhece-se que a aplicação de IAG nos processos licitatórios encontra-se em estágio inicial de desenvolvimento, demandando investigações empíricas complementares para validar a eficácia dos requisitos e técnicas identificados neste estudo conceitual. O estabelecimento destes fundamentos teóricos constitui, contudo, base necessária para o desenvolvimento dos *prompts* customizados e sua posterior validação.

REFERÊNCIAS

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, [2024]. Acesso em: 10 mar. 2025. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm.

_____. **Lei nº 14.133, de 1º de abril de 2021**. Lei de Licitações e Contratos Administrativos. Brasília, DF: Presidência da República, 2021. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14133.htm. Acesso em: 10 mar. 2025.

CONTINI, F. Unboxing Generative AI for the Legal Professions: Functions, Impacts and Governance. **International Journal for Court Administration**, v. 15, n. 2, p. 1-22, 2024. Disponível em: https://iacajournal.org/en/articles/10.36745/ijca.604. Acesso em: 28 set. 2024.

FAGBOHUN, O.; HARRISON, R. M.; DEREVENTSOV, A. **A Categorization of Prompting Techniques and Approaches for Large Language Models**: A Practitioner's Guide. Ophiuchus LLC, 2024. Disponível em: https://arxiv.org/pdf/2402.14837. Acesso em: 09 mar. 2025.

GAO, A. K. ***Prompt* Engineering for Large Language Models**: A Brief Guide with Examples for Non-Technical Readers. Stanford University, 2023. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4504303. Acesso em: 09 mar. 2025.

GROSSMAN, M.; GRIMM, S.; BROWN, A. Generative AI Policy for the Justice System. **Regent University Law Review**, v. 36, n. 1, p. 8-10, 2023.

HESSEL, R.; LEMES, L. **Perspectivas do uso da inteligência artificial generativa no Poder Judiciário brasileiro**. 2023. Disponível em: https://www.conjur.com.br/dl/artigo-ia-generativa-judiciario.pdf. Acesso em: 01 out. 2024.

IBM. **What is temperature in AI?** 2024. Acesso em: 13 jun. 2025. Disponível em: https://www.ibm.com/topics/temperature-ai.

JOVANOVIC, M.; CAMPBELL, M. Generative Artificial Intelligence: Trends and Prospects. **Computer Society**, v. 55, p. 107-120, 2022. Disponível em:

https://www.computer.org/csdl/magazine/co/2022/10/09903869/1H0G6xvtREk. Acesso em: 01 out. 2024.

KISSINGER, H.; SCHMIDT, E.; HUTTENLOCHER, D. **The Age of AI:** And Our Human Future. New York: Little, Brown and Company, 2021.

LEWIS, P.; PEREZ, E.; PIKTUS, A.; PETRONI, F.; KARPUKHIN, V.; GOYAL, N.; KÜTTLER, H.; LEWIS, M.; YIH, W.; ROCKTÄSCHEL, T.; RIEDEL, S.; KIELA, D. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In: **Proceedings of NeurIPS 2020**. 2020. Acesso em: 13 jun. 2025. Disponível em: https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html.

LIU, P.; YU, J.; JIANG, Z.; FU, H.; NEUBIG, G.; SHI, Z.; WANG, J. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. **ACM Computing Surveys,** v. 55, n. 9, p. 1-35, 2023. Acesso em: 07 out. 2024. Disponível em: https://dl.acm.org/doi/abs/10.1145/3560815.

MADAAN, A.; TANDON, N.; GUPTA, P.; HALLINAN, S.; GONG, L.; YANG, Y.; SCARIA, R.; PARMAR, H.; MISRA, B.; KALYAN, A.; SANTHANAM, S.; SHAH, S.; RAN, T.; MUDIGERE, D.; HALL, J.; CLARK, C.; WELLECK, S.; SABHARWAL, A. **Self-Refine**: Iterative Refinement with Self-Feedback. 2023. Acesso em: 13 jun. 2025. Disponível em: https://arxiv.org/abs/2303.17651.

PEEPERKORN, M.; VAN LEEUWEN, T.; BEULS, K. **Is Temperature the Creativity Parameter of Large Language Models?** 2024. Acesso em: 13 jun. 2025. Disponível em: https://arxiv.org/abs/2405.00492.

RAWTE, V.; SHETH, A.; DAS, A. **A survey of hallucination in large foundation models**. 2023. Disponível em: https://arxiv.org/abs/2309.05922. Acesso em: 01 out. 2024.

ROSSETTI, R.; GARCIA, K. Inteligência artificial generativa: questões jurídicas e éticas em torno do ChatGPT. **VirtuaJus**, Belo Horizonte, v. 8, n. 15, p. 253-264, 2023. Disponível em: https://periodicos.pucminas.br/index.php/virtuajus/article/download/30769/21848. Acesso em: 01 out. 2024.

SAHOO, P.; SINGH, A. K.; SAHA, S.; JAIN, V.; MONDAL, S.; CHADHA, A. **A Systematic Survey of *Prompt* Engineering in Large Language Models**: Techniques and Applications. Disponível em: arXiv preprint arXiv:2402.07927. Acesso em: 04 out. 2024.

SCHULHOFF, S.; ILIE, M., BALEPUR, N.; KAHADZE, K., LIU, A.; SI, C.; BHAT, Y. K.; DENG, C.; DONG, H.; DONG, Z.; ELARABY, M.; FUNG, Y.; GALLEY, M.; GAO, T.; GERO, K.; HASHEMI, H.; HEINEMANN, B.; KAZEMNEJAD, A.; LAHLOU, S.; LAI, Y.; LEE, J.; LIU, C. Y.; LIU, N. F.; MADDELA, M.; MAMERTINE, K.; MENGYANG, C.; MIALON, G.; MIRESHGHALLAH, F.; MORTON, A. S.; NATARAJAN, N.; NEUBIG, G.; PASUPAT, P.; RAFFEL, C.; RAJA, K.; ROCKTÄSCHEL, T.; RUD, T.; SETH, S.; SHI, T.; SHINAVER, J.; STEWART, I.; TAN, S.; THRUSH, T.; TRAN, V.; VYAS, H.; VU, T.; WAN, J.; WANG, A.; WANG, W.; WEST, P.; YU, Q.; ZHANG, D.; ZHENG, J.; ZIEGLER, D. A.; RESNIK, P. (2024). **The *prompt* report**: A systematic survey of prompting techniques. Disponível em: https://arxiv.org/abs/2406.06608. Acesso em 9 mar. 2025.

SHEN, Y. KANG, K.; WANG, Y.; ZHANG, H.; WANG, W.; CHEN, W.; ZHANG, Y.; HAN, Z.; LI, Y.; GU, J.; YANG, W.; DING, Y.; ZOU, L.; ZOU, X.; SHOU, L.; ZHOU, M.; KIM, D. **Large language models:** A Survey. 2023. Acesso em: 01 out. 2024. Disponível em: https://arxiv.org/abs/2303.18223.

TAKALE, D. G.; MAHALLE, P. N.; SULE, B. Advancements and Applications of Generative Artificial Intelligence. **Journal of Information Technology and Sciences**, v. 10, n. 1, p. 20-27, 2024. Disponível em: https://www.researchgate.net/profile/Dattatray-Takale/publication/378942763_Advancements_and_Applications_of_Generative_Artificial_Intelligence/links/65f2d4d632321b2cff78c88d/Advancements-and-Applications-of-Generative-Artificial-Intelligence.pdf. Acesso em: 01 out. 2024.

TONMOY, S. M. FAISAL, F.; RAHMAN, T.; WANG, H.; ANANIADOU, S.; ALI, J.; FIROOZI, R.; MOHIUDDIN, T.; WAHAB, M.; VAFA, P. F. **A comprehensive survey of hallucination mitigation techniques in large language models**. 2024. Disponível em: https://arxiv.org/abs/2401.01313. Acesso em: 01 out. 2024.

TORRES, R. C. L. **Leis de Licitações públicas comentadas**. 12. ed. rev., ampl. e atual. São Paulo: Ed. Juspodiym, 2021.

VASWANI, A. SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is All You Need. In: Advances In Neural Information Processing Systems, 30, 2017. **Proceedings** [...]. 2017. p. 5998-6008. Disponível em: https://www.codetds.com/article/15517262. Acesso em: 28 set. 2023.

XU, Z.; JAIN, S.; KANKANHALLI, M. **Hallucination is Inevitable**: An Innate Limitation of Large Language Models. 2024. Disponível em: https://arxiv.org/abs/2401.11817. Acesso em: 28 set. 2024.

YAO, S.; ZHAO, J.; YU, D.; DU, N.; SHAFRAN, I.; NARASIMHAN, K.; CAO, Y. **ReAct**: Synergizing Reasoning and Acting in Language Models. 2022. Disponível em: https://arxiv.org/abs/2210.03629. Acesso em: 13 jun. 2025.

4. INTELIGÊNCIA ARTIFICIAL GENERATIVA NAS LICITAÇÕES PÚBLICAS: UM ESTUDO DE CASO NO TRIBUNAL DE JUSTIÇA DE PERNAMBUCO

4.1. INTRODUÇÃO

A transformação digital do Poder Judiciário Brasileiro, impulsionada por iniciativas como o Programa Justiça 4.0, tem fomentado a busca por soluções tecnológicas capazes de otimizar processos administrativos e jurisdicionais. Nesse contexto, os procedimentos licitatórios representam uma área crítica da administração judiciária, caracterizada por sua complexidade normativa e pelo alto impacto na eficiência operacional das instituições. A Inteligência Artificial Generativa (IAG), ao possibilitar a automação na criação de documentos complexos, emerge como uma ferramenta promissora para modernizar e agilizar esses processos.

O TJPE apresenta-se como um campo fértil para a implementação experimental de soluções baseadas em IAG. A modalidade de Dispensa de Licitação, prevista no artigo 75 da Lei nº 14.133/2021, configura-se como um ponto de entrada estratégico para essa inovação, considerando seu volume expressivo e a relativa padronização de seus requisitos documentais.

No entanto, a eficácia da IAG na geração de artefatos licitatórios depende, fundamentalmente, da qualidade das instruções fornecidas aos modelos de linguagem – os *prompts* . Conforme discutido por Sahoo *et al* . (2024), a engenharia de *prompts* representa um campo emergente que demanda investigação científica rigorosa, especialmente quando aplicada a domínios altamente regulamentados como o das contratações públicas.

Outro aspecto crítico refere-se ao fenômeno das alucinações -- geração de conteúdo factualmente incorreto ou inconsistente -- que impõe desafios à implementação segura e confiável de IAG no contexto jurídico-administrativo. A pesquisa do CNJ (2024e) revelou que 48,9% dos magistrados relataram "falta de confiança nos resultados gerados" por ferramentas de IAG, evidenciando a necessidade de estratégias específicas de mitigação desses riscos.

Diante desse cenário, surge a seguinte questão de pesquisa: de que modo a engenharia de *prompts* pode ser estruturada para viabilizar o uso seguro e transparente da IAG na elaboração de artefatos administrativos voltados à contratação direta por Dispensa de Licitação, no âmbito do TJPE?

O estudo tem como objetivo geral implementar e validar empiricamente um conjunto de *prompts* customizados para geração de Termos de Referência na modalidade de Dispensa de Licitação, contemplando os aspectos éticos e de transparência exigidos pela legislação vigente e pelas normativas do CNJ. Como objetivos específicos, buscou-se: (i) desenvolver *prompts* estruturados segundo metodologias avançadas de engenharia; (ii) testar e validar os *prompts* através de experimentação controlada; (iii) avaliar a precisão, consistência e incidência de alucinações dos artefatos gerados; e (iv) analisar os impactos de eficiência temporal e qualidade comparativa.

A relevância desta investigação amplifica-se diante dos desafios identificados na literatura especializada sobre aplicação de IA no judiciário brasileiro. A escolha do Termo de Referência como objeto central desta pesquisa justifica-se por sua importância técnica nos processos de Dispensa de Licitação e sua estrutura padronizada, definida pelos dez elementos obrigatórios do artigo 6º, inciso XXIII da Lei nº 14.133/2021, que facilita o desenvolvimento de *prompts* específicos.

O capítulo está estruturado em seis seções complementares: a metodologia da pesquisa (4.2) apresenta o delineamento experimental e os procedimentos de validação; o desenvolvimento dos *prompts* (4.3) detalha a arquitetura híbrida implementada; a experimentação e coleta de dados (4.4) descreve os procedimentos empíricos; os resultados e análise (4.5) consolidam os achados quantitativos; e a discussão (4.6) interpreta os resultados à luz da literatura e das diretrizes de governança. Por fim, as considerações finais (4.7) sintetizam as contribuições e estabelecem conexões com o produto técnico-tecnológico desenvolvido.

4.2. METODOLOGIA DA PESQUISA

A metodologia científica constitui o alicerce estrutural que orienta o processo investigativo, definindo os procedimentos, técnicas e instrumentos que permitirão alcançar os objetivos propostos de forma sistemática e coerente (Marconi; Lakatos, 2021). Além disso, "para que um conhecimento possa ser considerado científico, torna-se necessário identificar as operações mentais e técnicas que possibilitam a sua verificação" (Gil, 2022, p. 8). Assim, este capítulo apresenta o percurso metodológico adotado para viabilizar a integração da IAG nos processos licitatórios do Poder Judiciário brasileiro.

4.2.1. Classificação

Esta investigação caracteriza-se, quanto à natureza, como uma pesquisa aplicada, pois objetiva "gerar conhecimentos para aplicação prática dirigidos à solução de problemas

específicos" (Prodanov; Freitas, 2013). Segundo Thiollent (2011), a pesquisa aplicada busca elaborar diagnósticos, identificar problemas e propor soluções, características que se alinham perfeitamente ao presente estudo, que visa desenvolver e implementar soluções baseadas em IAG para otimizar os processos licitatórios no âmbito do Poder Judiciário.

Quanto aos objetivos, trata-se de uma pesquisa exploratório-descritiva. O caráter exploratório justifica-se pela busca de maior familiaridade com o fenômeno estudado (Gil, 2022), especificamente a aplicação de IAG em licitações públicas, tema ainda emergente na literatura científica brasileira. Simultaneamente, classifica-se como descritiva por buscar caracterizar as particularidades do processo licitatório e estabelecer relações entre as variáveis envolvidas na implementação da tecnologia generativa (Vergara, 2016).

No que concerne à abordagem, a pesquisa adota métodos mistos (quali-quantitativos). A dimensão qualitativa manifesta-se na compreensão aprofundada das especificidades normativas e administrativas dos processos de Dispensa de Licitação, bem como na análise interpretativa da adequação dos *prompts* desenvolvidos aos requisitos éticos e de transparência. A dimensão quantitativa evidencia-se nas métricas estruturadas de avaliação empregadas pelos especialistas, incluindo *scores* de conformidade legal, indicadores de eficiência temporal e percentuais de adequação técnica (Creswell; Creswell, 2021).

Quanto aos procedimentos técnicos, a pesquisa emprega uma combinação metodológica que inclui:

1. Pesquisa bibliográfica para fundamentação teórica sobre engenharia de *prompts* e marcos normativos; pesquisa documental para análise da legislação vigente e normativas do CNJ;
2. Estudo de caso focado na realidade específica do TJPE; e
3. Pesquisa experimental para desenvolvimento, teste e validação controlada dos *prompts* customizados. Segundo Severino (2017), o procedimento experimental caracteriza-se pela manipulação direta das variáveis relacionadas ao objeto de estudo, o que se materializa no desenvolvimento iterativo e teste sistemático dos *prompts* de IAG.

Esta pesquisa alinha-se ao paradigma do *Design Science Research* (DSR), abordagem metodológica que se concentra na construção e avaliação de artefatos destinados à resolução de problemas práticos identificados (Dresch; Lacerda; Antunes, 2015). No contexto desta investigação, os artefatos desenvolvidos consistem no sistema de *prompts* estruturados e nas diretrizes de implementação, ambos validados através de aplicação empírica no ambiente institucional do TJPE.

A natureza interdisciplinar desta pesquisa, que articula conhecimentos de Direito Administrativo, Ciência da Computação e Administração Pública, reflete a complexidade inerente à implementação responsável de tecnologias emergentes no setor público. Esta convergência disciplinar é essencial para assegurar que as soluções desenvolvidas atendam simultaneamente aos critérios de eficiência operacional, conformidade normativa e responsabilidade ética.

Por fim, a incorporação explícita de aspectos éticos e de transparência, alinhada às diretrizes estabelecidas pela Resolução CNJ nº 615/2025, confere à pesquisa uma dimensão de responsabilidade social que transcende os objetivos puramente técnicos, contribuindo para o desenvolvimento de práticas de governança algorítmica no setor público brasileiro.

4.2.2. Procedimentos Metodológicos

O desenvolvimento desta pesquisa estruturou-se em quatro etapas sequenciais e complementares, organizadas para garantir rigor científico e aplicabilidade prática dos resultados. A metodologia adotada combina elementos de pesquisa experimental com desenvolvimento tecnológico aplicado, seguindo as diretrizes de Thiollent (2011) para pesquisas que visam à geração de soluções práticas para problemas específicos.

A arquitetura metodológica fundamenta-se na integração entre teoria e prática, característica essencial do mestrado profissional, que demanda não apenas a produção de conhecimento científico, mas sua aplicação efetiva na transformação da realidade organizacional. Esta abordagem permite que cada etapa da pesquisa contribua tanto para o avanço do conhecimento quanto para a solução dos problemas identificados no contexto do TJPE.

4.2.2.1. Primeira Etapa: Fundamentação Teórica e Normativa

A primeira etapa concentrou-se na construção de alicerces teóricos e normativos sólidos, através de revisão sistemática da literatura e mapeamento do marco regulatório. Esta fase teve como objetivo identificar requisitos técnicos e jurídicos essenciais para desenvolvimento de soluções automatizadas no contexto administrativo.

O levantamento bibliográfico abrangeu três dimensões complementares: literatura internacional sobre engenharia de *prompts* e aplicações de IAG em administração pública; estudos nacionais sobre implementação de IA no Poder Judiciário brasileiro; e pesquisas específicas sobre mitigação de riscos em sistemas automatizados de tomada de decisão. A seleção das fontes seguiu critérios de relevância, atualidade e qualidade metodológica,

priorizando publicações dos últimos cinco anos em periódicos indexados e documentos oficiais de organismos internacionais.

A análise documental contemplou o exame sistematizado da legislação aplicável, incluindo a Constituição Federal de 1988 (princípios da administração pública), a Lei nº 14.133/2021 (Nova Lei de Licitações), as Resoluções do CNJ nº 332/2020 e nº 615/2025 (uso de IA no Judiciário), e normativos internos do TJPE. Esta análise permitiu identificar não apenas os requisitos legais explícitos, mas também as implicações normativas implícitas para implementação de tecnologias automatizadas em processos administrativos.

O mapeamento de experiências envolveu a sistematização de casos de uso de IAG em organizações públicas nacionais, com foco nas iniciativas piloto implementadas por tribunais brasileiros. Esta análise comparativa forneceu *insights* valiosos sobre boas práticas, limitações identificadas e estratégias de mitigação de riscos.

A primeira etapa concentrou-se na fundamentação teórica e normativa, por meio de revisão sistemática da literatura sobre IAG aplicada à administração pública e mapeamento exaustivo do marco regulatório das licitações brasileiras. Esta fase visou identificar requisitos técnicos e jurídicos essenciais para desenvolvimento de soluções automatizadas no contexto administrativo.

4.2.2.2. Segunda Etapa: Seleção e Adaptação de Técnicas de Engenharia de *Prompts*

Com base nos achados da revisão de literatura e considerando as especificidades do domínio licitatório, procedeu-se à seleção criteriosa de técnicas avançadas de engenharia de *prompts*. Esta etapa envolveu análise comparativa de diferentes abordagens metodológicas e sua adequação aos requisitos de transparência, precisão e conformidade legal exigidos pela legislação vigente.

A análise técnica das metodologias disponíveis baseou-se em critérios objetivos: capacidade de estruturação de raciocínio lógico-jurídico; eficácia na mitigação de alucinações; compatibilidade com bases de dados externas; facilidade de implementação e manutenção; e aderência aos princípios de transparência e auditabilidade estabelecidos pelo CNJ.

Foram inicialmente selecionadas cinco técnicas principais: ***Chain-of-Thought*** **(CoT)** para estruturação do raciocínio jurídico-administrativo; ***Retrieval-Augmented Generation*** **(RAG)** para consulta a bases normativas atualizadas*;* ***ReAct*** para refinamento iterativo; ***Self-Refine*** para controle de qualidade automatizado; *e* ***Few-Shot Learning*** para padronização institucional. Durante os testes preliminares, identificou-se que a técnica *ReAct* apresentava

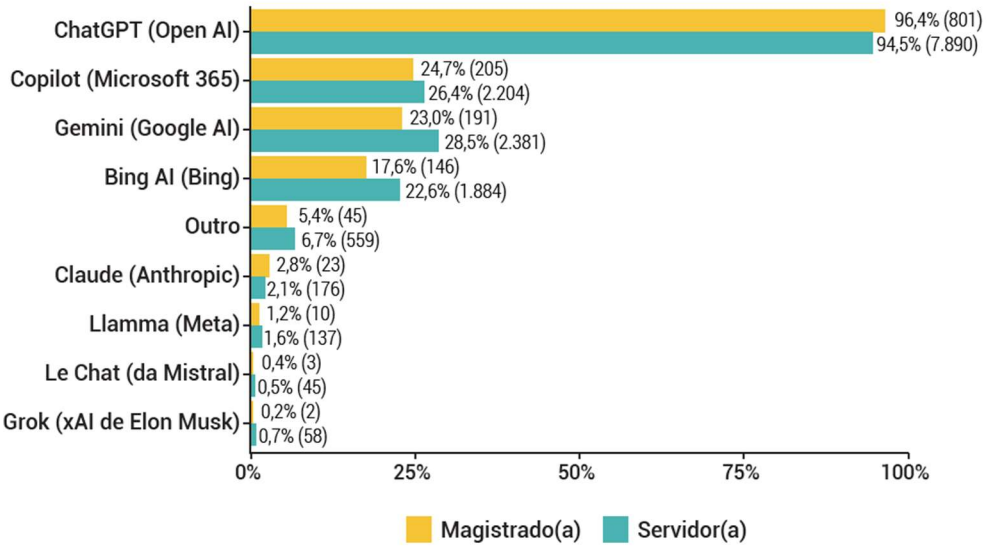
limitações significativas, sendo substituída por *Few-Shot Learning* conforme detalhado na seção 4.2.2.3.1.

A configuração de parâmetros incluiu a definição de Persona (analista jurídico especializado em licitações do TJPE) e Temperatura baixa (≤0.2) para garantir precisão e conformidade normativa. Esta configuração fundamentou-se tanto nas evidências empíricas da literatura quanto nas especificidades do contexto administrativo, que exige previsibilidade e aderência rigorosa aos marcos normativos.

4.2.2.3. Terceira Etapa: Seleção de Ferramentas de IAG

A terceira etapa compreendeu a justificativa técnica para escolha da ferramenta de IAG mais adequada aos objetivos da pesquisa. Como demonstra a figura 24, considerou-se o ChatGPT devido à sua predominância no contexto judiciário brasileiro, conforme evidenciado pelo relatório do CNJ (2024e), que demonstra utilização por 96,4% dos magistrados e 94,5% dos servidores que já experimentaram tecnologias de IAG.

Figura 24: IAGs utilizadas no Poder Judiciário Brasileiro



Fonte: CNJ (2024e , p. 53)

A avaliação comparativa seguiu critérios técnicos objetivos: capacidade de processamento de documentos complexos; eficácia em *Few-Shot Learning;* estabilidade na geração de documentos completos; compatibilidade com as técnicas selecionadas; e adequação aos requisitos de transparência e auditabilidade.

4.2.2.3.1. Limitações identificadas no ChatGPT

Durante a fase de testes preliminares, identificaram-se limitações críticas que comprometeram a aplicabilidade do ChatGPT para os objetivos desta pesquisa. Especificamente para a técnica *Few-Shot Learning* - essencial para reproduzir padrões documentais específicos do TJPE - o ChatGPT apresentou desempenho inadequado em duas dimensões principais:

- Geração sistemática de documentos incompletos, substituindo seções estruturadas por expressões genéricas como "conforme modelo" ou "seguindo padrão estabelecido", mesmo quando fornecidos exemplos detalhados e instruções específicas para completude documental. A figura 25 exemplifica essa limitação, demonstrando como o ChatGPT sistematicamente omitia seções do Termo de Referência, comprometendo a aplicabilidade prática dos documentos gerados;
- Dificuldade em processar adequadamente os exemplos fornecidos na técnica *Few-Shot*, conforme identificado por Liu *et al*. (2023) e evidenciado por Brown *et al.* (2020), que demonstraram como diferentes arquiteturas de LLM apresentam variações significativas na eficácia de *few-shot learning*, especialmente em tarefas que exigem reprodução fiel de formatos estruturados complexos.

Figura 25: Exemplo de limitação do ChatGPT na geração completa de Termos de Referência



Fonte: Elaborado pela autora (2025)

4.2.2.3.2. Transição para o Claude

Diante dessas limitações, procedeu-se à avaliação de ferramentas alternativas, culminando na seleção do Claude, desenvolvido pela Anthropic[6]. Esta escolha fundamentou-se em três critérios principais: Capacidade superior em *Few-Shot Learning,* conforme evidenciado em testes comparativos, os quais demonstraram que o Claude processa mais eficazmente exemplos estruturados, reproduzindo fielmente formatos complexos sem recurso a simplificações ou omissões (Anthropic, 2024); geração consistente de documentos completos sem omissões ou simplificações inadequadas; e alinhamento com princípios de *Constitutional AI* (Bai et al., 2022), particularmente adequado para contextos que exigem conformidade normativa e transparência.

A mudança de ferramenta, embora não prevista inicialmente, demonstrou-se necessária para garantir a viabilidade técnica da pesquisa e a qualidade dos artefatos gerados, mantendo-se alinhada aos princípios de rigor científico e aplicabilidade prática que orientam este estudo.

4.2.2.4. Quarta Etapa: Desenvolvimento e Validação Experimental

A etapa final materializou-se no desenvolvimento iterativo e validação experimental dos *prompts* estruturados, seguindo princípios de prototipação rápida e teste empírico (Nielsen, 1993). Esta fase incluiu múltiplos ciclos de refinamento baseados em *feedback* de especialistas e validação através de casos reais do TJPE.

Desenvolvimento Interativo

O desenvolvimento iterativo seguiu a metodologia DSR, conforme estabelecido por Dresch, Lacerda e Antunes (2015). O DSR caracteriza-se pela construção e avaliação de artefatos destinados à resolução de problemas práticos, através de ciclos iterativos que incluem: identificação do problema, definição dos objetivos da solução, *design* e desenvolvimento do artefato, demonstração, avaliação e comunicação dos resultados. Cada ciclo incluía prototipação, teste, avaliação por especialistas e refinamento, até atingir estabilidade funcional e conformidade com todos os requisitos estabelecidos.

Durante esta etapa, identificou-se a necessidade de substituir a técnica *ReAct por Few-Shot Learning*. A técnica *ReAct* apresentava limitações significativas: geração de documentos excessivamente verbosos com redundâncias desnecessárias; inadequação para a estrutura linear e sequencial dos Termos de Referência; e fragmentação da lógica argumentativa. O *Few-Shot*

[6] Anthropic é uma empresa de pesquisa em segurança de IA fundada em 2021, especializada no desenvolvimento de sistemas de inteligência artificial seguros e alinhados com valores humanos através de técnicas como "*Constitutional AI*" (Anthropic, 2024; Bai et al., 2022)

Learning demonstrou capacidade superior de reproduzir padrões específicos através de exemplos orientadores, mantendo concisão e objetividade requeridas em documentos administrativos.

Validação Experimental

A validação experimental consistiu na aplicação do sistema em cinco casos reais de Dispensa de Licitação do TJPE, representativos das contratações realizadas e concluídas entre 2024 e 2025. Cada documento original foi pareado com uma versão automatizada, permitindo análise comparativa rigorosa. A avaliação foi conduzida por três especialistas utilizando instrumento estruturado baseado nos dez elementos obrigatórios estabelecidos pela Lei nº 14.133/2021.

Critérios de Validação

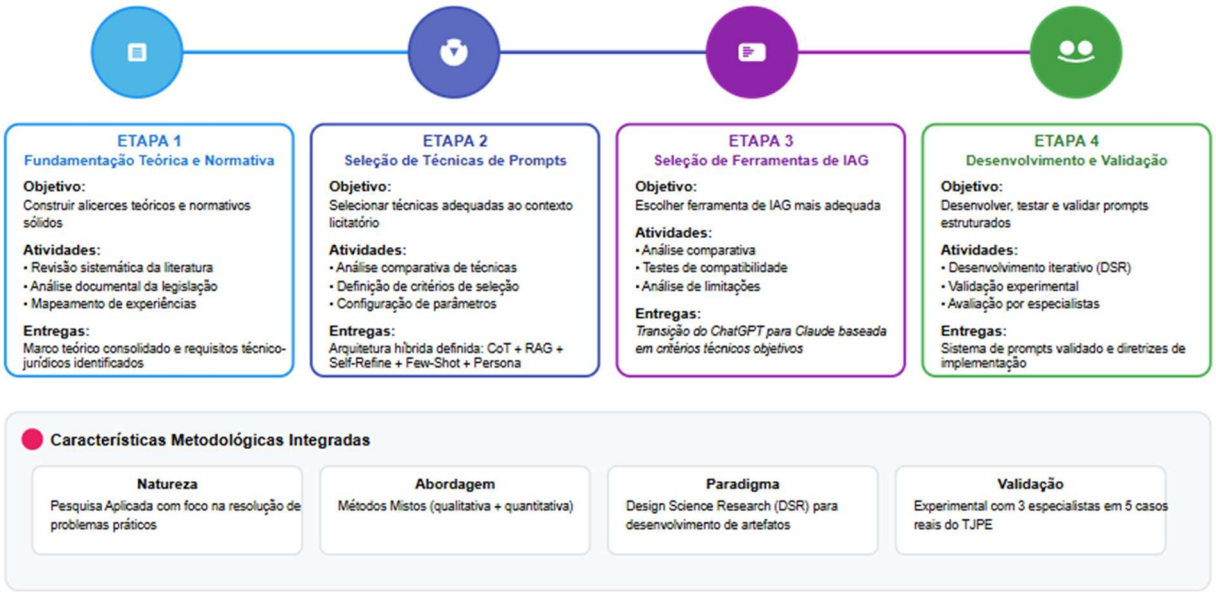
Os critérios de validação contemplaram três dimensões: conformidade legal (aderência aos requisitos normativos); qualidade técnica (precisão, completude e adequação); e adequação institucional (alinhamento aos padrões do TJPE). Esta abordagem multidimensional assegurou avaliação abrangente da solução desenvolvida, considerando tanto aspectos técnicos quanto sua aplicabilidade prática no contexto organizacional específico.

A metodologia adotada nesta etapa garantiu que o produto final atendesse simultaneamente aos critérios de rigor científico, aplicabilidade prática e conformidade às exigências normativas do setor público, estabelecendo as bases para futuras replicações e expansões da pesquisa.

A convergência metodológica entre as quatro etapas descritas fundamenta-se na necessidade de articular rigor científico com aplicabilidade prática, característica essencial do mestrado profissional. Esta arquitetura metodológica, estruturada de forma sequencial, mas com elementos de retroalimentação, pode ser visualizada de forma integrada na figura 26, que sintetiza o percurso completo da investigação.

Como evidenciado na figura 26, a metodologia adotada transcende a mera aplicação de técnicas isoladas, constituindo um sistema integrado que assegura tanto a validade científica quanto a viabilidade prática da solução desenvolvida. Esta arquitetura metodológica estabelece as bases para compreender não apenas os resultados obtidos, mas também a replicabilidade e adaptabilidade da solução proposta a outros contextos institucionais, aspecto fundamental para a contribuição científica e prática desta investigação.

Figura 26: *Design* Metodológico - Desenvolvimento de *Prompts* para IAG em Licitações Públicas



Fonte: Elaborado pela autora (2025)

4.2.3. Limitações Metodológicas

Reconhecem-se como limitações metodológicas deste estudo aspectos que podem impactar a generalização dos resultados e a robustez das conclusões, embora tenham sido adotadas estratégias específicas de mitigação para cada uma delas.

A primeira limitação refere-se ao número restrito de avaliadores (três especialistas), que pode limitar a representatividade das perspectivas e reduzir a robustez estatística das análises comparativas. Esta limitação é inerente ao contexto do TJPE, onde o número de servidores especializados em licitações é naturalmente reduzido, e foi mitigada com a seleção criteriosa de avaliadores com *expertise* complementar e experiência diversificada em licitações públicas, incluindo diferentes perfis profissionais (pregoeiro, gestor de contratos e especialista em TI), visando ampliar a representatividade das perspectivas.

A segunda limitação decorre da rápida evolução das ferramentas de IAG, que pode tornar alguns achados temporalmente situados. As tecnologias de IAG experimentam atualizações frequentes que podem alterar significativamente suas capacidades e limitações. Para atenuar este aspecto, foi realizada documentação detalhada dos procedimentos metodológicos, permitindo sua replicação e atualização, além de terem sido estabelecidos critérios de avaliação baseados em fundamentos técnicos duradouros, como conformidade normativa e transparência, que permanecem relevantes independentemente das evoluções tecnológicas específicas.

A terceira limitação relaciona-se à complexidade e variedade dos processos licitatórios nos diferentes órgãos do Judiciário, o que pode restringir a generalização dos resultados. Cada tribunal possui particularidades organizacionais, volume de contratações e especificidades técnicas que podem influenciar a aplicabilidade da solução desenvolvida. Esta limitação foi endereçada através do desenvolvimento de *prompts* baseados na legislação federal (Lei nº 14.133/2021) e nas diretrizes do CNJ, que constituem marcos normativos aplicáveis a todo o Poder Judiciário, além da incorporação de elementos de flexibilidade e customização que permitem adaptação a diferentes contextos institucionais.

A quarta limitação refere-se à necessidade de avaliação prolongada para verificar a robustez da solução ao longo do tempo e em diferentes contextos operacionais. A pesquisa concentrou-se em um período específico e em casos selecionados, não contemplando variações sazonais, mudanças normativas futuras ou diferentes perfis de contratação. Para mitigar esta limitação, foram estabelecidos protocolos de monitoramento contínuo e mecanismos de atualização dos *prompts*, além de diretrizes para implementação gradual e acompanhamento sistemático dos resultados.

A quinta limitação diz respeito ao escopo institucional específico do TJPE, que possui características organizacionais, volume processual e práticas administrativas particulares. Esta delimitação, embora necessária para viabilizar a pesquisa experimental, pode restringir a aplicabilidade imediata dos achados em outros contextos. A estratégia de mitigação incluiu a triangulação de fontes e métodos de coleta de dados, incluindo análise documental, experimentação prática e validação por especialistas, além da transparência metodológica quanto aos procedimentos adotados e às especificidades do contexto estudado.

Adicionalmente, reconhece-se como limitação a participação dual da pesquisadora como avaliadora, o que pode introduzir vieses inconscientes na análise. Esta situação foi endereçada através da implementação de mecanismos de controle de viés, incluindo o desenvolvimento de instrumento de avaliação estruturado (*checklist*) baseado em critérios objetivos da Lei nº 14.133/2021, reduzindo a subjetividade das análises, e pela transparência metodológica explícita sobre este papel dual, permitindo que futuros pesquisadores considerem adequadamente esta variável na interpretação dos resultados.

Outro aspecto importante refere-se às restrições temporais, que impactaram os procedimentos de coleta de dados. Devido ao prazo exíguo da pesquisa experimental, optou-se por gerar previamente os cinco Termos de Referência automatizados via aplicação controlada do sistema de *prompts* desenvolvido, utilizando como base os documentos preparatórios (DFD

e ETP) dos processos selecionados. Os avaliadores receberam, portanto, documentos já processados pelo sistema, em vez de acompanharem o processo de geração em tempo real.

Esta abordagem, embora preserve a validade da comparação entre documentos manuais e automatizados, limita a capacidade de avaliar aspectos operacionais como tempo de interação usuário-sistema e eventuais necessidades de refinamento iterativo durante o processo de geração. Para mitigar esta limitação, foi documentado detalhadamente o processo de geração de cada TR, incluindo dados de entrada utilizados, configurações aplicadas e tempo de processamento, assegurando replicabilidade e transparência metodológica.

Finalmente, destaca-se que o caráter pioneiro da aplicação de engenharia de *prompts* em processos licitatórios no contexto brasileiro constitui simultaneamente uma contribuição e uma limitação, na medida em que não existem parâmetros consolidados de comparação ou *benchmarks* estabelecidos na literatura nacional. Esta limitação foi abordada por meio da construção de referencial teórico sólido baseado na literatura internacional e na adaptação criteriosa de metodologias consolidadas para o contexto específico da administração pública brasileira.

Apesar dessas limitações, a metodologia adotada assegura que os resultados obtidos ofereçam contribuições válidas e relevantes para o avanço do conhecimento científico sobre aplicação de IAG na administração pública, estabelecendo bases sólidas para futuras investigações e replicações em diferentes contextos organizacionais.

4.3. DESENVOLVIMENTO DOS *PROMPTS*

O desenvolvimento dos *prompts* constituiu processo iterativo de *design*, teste e refinamento, fundamentado na metodologia DSR e orientado pelos requisitos técnico-jurídicos identificados na fase de fundamentação teórica. Este processo materializou-se na construção de uma arquitetura híbrida que integra múltiplas técnicas de engenharia de *prompts*, customizada para atender às especificidades dos processos de Dispensa de Licitação do TJPE.

A implementação prática seguiu princípios de prototipação rápida e validação empírica (Nielsen, 1993), com ciclos sistemáticos de desenvolvimento que incluíram testes funcionais, avaliação por especialistas e refinamentos baseados em *feedback* operacional. O objetivo central foi traduzir os requisitos teóricos identificados na literatura e nos normativos jurídicos em instruções práticas capazes de orientar modelos de IAG na geração de documentos administrativos precisos, conformes e auditáveis.

4.3.1. Arquitetura Híbrida Implementada

A arquitetura final dos *prompts* integra cinco componentes técnicos principais, cada um atendendo a requisitos específicos do contexto licitatório. Esta configuração híbrida fundamenta-se na premissa de que a complexidade e criticidade dos documentos administrativos demandam abordagem multidimensional, que contemple simultaneamente precisão normativa, estruturação lógica, atualização factual e controle de qualidade.

O **CoT** foi implementado como estrutura base para organizar o raciocínio jurídico-administrativo, garantindo que os documentos sigam sequência lógica alinhada aos princípios de fundamentação e motivação dos atos administrativos. Na prática, esta técnica manifesta-se através de instruções específicas que orientam o modelo a explicitar cada etapa do processo decisório, desde a identificação da necessidade até a justificativa da modalidade de contratação escolhida.

O **RAG** foi configurado para consulta automática a bases normativas atualizadas, incluindo a Lei nº 14.133/2021, resoluções do CNJ e jurisprudência administrativa consolidada. Esta implementação assegura que os documentos gerados estejam fundamentados em fontes confiáveis e atualizadas, mitigando riscos de desatualização normativa e garantindo aderência ao ordenamento jurídico vigente.

O ***Self-Refine*** opera como mecanismo de controle de qualidade automatizado, instruindo o modelo a revisar sistematicamente suas próprias saídas antes da apresentação final. Esta técnica implementa verificações de consistência interna, conformidade normativa e adequação formal, reduzindo significativamente a incidência de inconsistências e imprecisões.

O ***Few-Shot Learning*** foi customizado com exemplos representativos de Termos de Referência do TJPE, permitindo que o modelo aprenda os padrões institucionais específicos através de demonstrações concretas. Esta abordagem garante aderência aos formatos, terminologias e níveis de detalhamento esperados pela instituição.

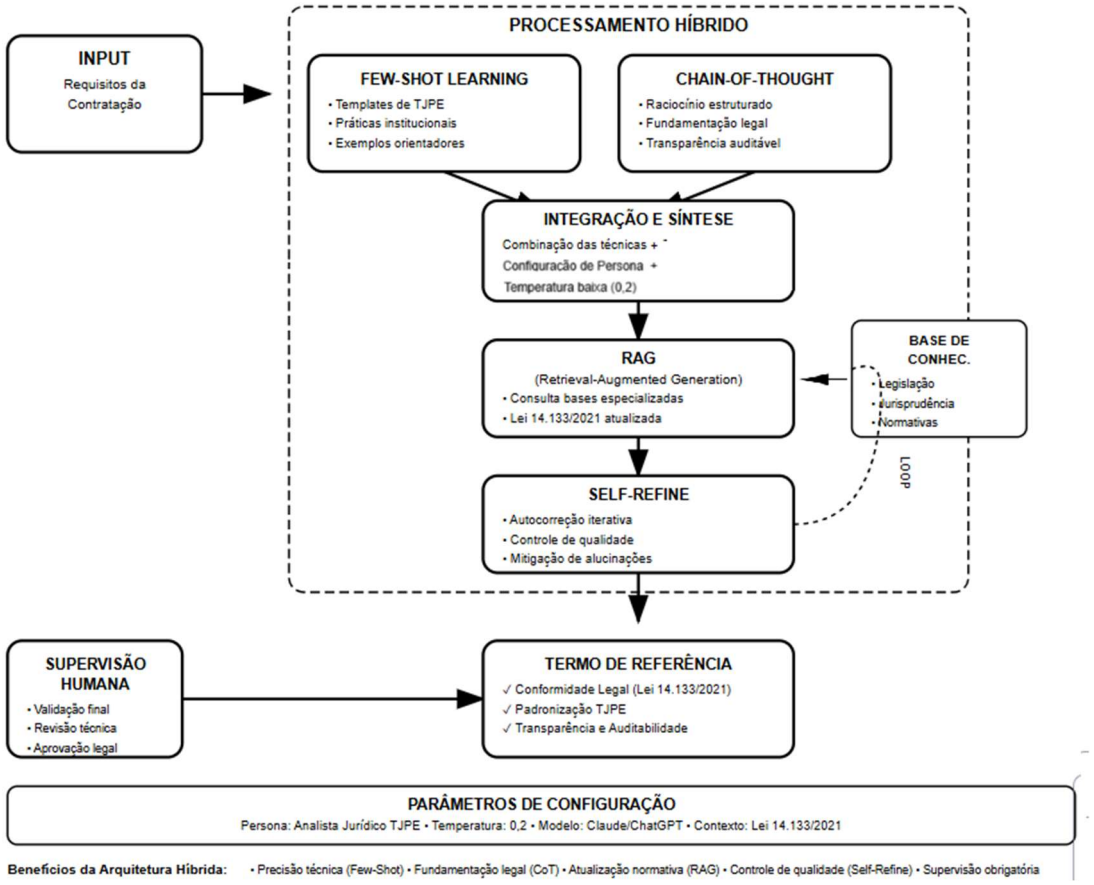
A **configuração de Persona** estabelece o perfil contextual de "analista jurídico especializado em licitações do TJPE", orientando estilo, terminologia e profundidade técnica adequados ao domínio específico. Complementarmente, a **Temperatura baixa** (0.2) assegura previsibilidade e precisão nas saídas, características essenciais para documentos que fundamentam decisões de gasto público.

A operacionalização desta arquitetura híbrida materializa-se através de um fluxo de processamento estruturado que integra sequencialmente as cinco técnicas selecionadas, desde a entrada dos requisitos de contratação até a geração do Termo de Referência final. A figura 27 ilustra esquematicamente esta arquitetura, evidenciando como cada componente contribui para

a construção progressiva do documento e os pontos de controle que asseguram qualidade e conformidade.

O processamento inicia-se com a entrada dos requisitos específicos da contratação (*input*), que alimentam simultaneamente as técnicas *Few-Shot Learning* e *Chain-of-Thought.* Esta abordagem paralela permite que o modelo acesse tanto os padrões institucionais (através dos exemplos do TJPE) quanto estruture o raciocínio jurídico-administrativo de forma transparente e auditável. A integração subsequente dessas técnicas, combinada com a configuração de Persona e Temperatura baixa, estabelece as bases contextuais para as etapas seguintes do processamento.

Figura 27: Arquitetura de *prompt* híbrido proposto



Fonte: Baseado em Schulhoff *et al*. (2024), Liu *et al*. (2023) e na metodologia da pesquisa (2025)

O módulo RAG opera como ponte entre o conhecimento interno do modelo e as bases normativas externas, consultando automaticamente a legislação vigente e a jurisprudência administrativa para fundamentar adequadamente cada seção do documento. Por fim, *o Self-Refine* implementa controle de qualidade automatizado, realizando verificações sistemáticas

antes da supervisão humana final, que permanece como elemento indispensável para validação e aprovação legal.

4.3.2. Processo de Refinamento Iterativo

O desenvolvimento seguiu metodologia iterativa estruturada em ciclos de prototipação, teste, avaliação e refinamento, cada um com duração média de uma semana. Este processo permitiu identificação e correção progressiva de limitações, ajuste fino dos parâmetros e otimização da qualidade dos documentos gerados:

- **Primeiro Ciclo - prototipação inicial:** desenvolvimento da versão base dos *prompts*, integrando as cinco técnicas selecionadas em estrutura coesa. Testes preliminares revelaram necessidade de maior especificação das instruções e melhor balanceamento entre as diferentes técnicas;
- **Segundo Ciclo - ajustes estruturais:** refinamento da arquitetura baseado nos achados do primeiro ciclo, com ênfase na melhor integração entre CoT e RAG. Identificou-se a necessidade de substituir a técnica ReAct por *Few-Shot Learning*, conforme detalhado na seção metodológica;
- **Terceiro Ciclo - otimização contextual:** *customização* dos *prompts* para o contexto específico do TJPE, incorporando terminologias, formatos e padrões institucionais. Este ciclo incluiu validação preliminar com servidores especializados;
- **Quarto Ciclo - validação final:** teste da versão otimizada em casos reais, com avaliação sistemática por três especialistas. Os ajustes finais focaram na melhoria da consistência e na redução de variações desnecessárias.

Cada ciclo incluiu métricas de avaliação específicas: conformidade legal (percentual de aderência aos requisitos da Lei nº 14.133/2021), completude documental (presença de todos os elementos obrigatórios), e adequação institucional (alinhamento aos padrões do TJPE). A convergência dessas métricas em patamares satisfatórios determinou a conclusão do processo iterativo.

4.3.3. Estrutura Final dos *Prompts* Validados

A versão final dos *prompts* estrutura-se em quatro blocos funcionais integrados: contextualização, instruções técnicas, exemplos orientadores e parâmetros de controle. Esta organização assegura clareza instrucional e facilita manutenção e atualizações futuras.

O **Bloco de Contextualização** estabelece a persona do analista jurídico e define o contexto operacional (geração de Termos de Referência para dispensas de licitação no TJPE). Esta seção inclui instruções sobre tom, estilo e nível de formalidade esperados.

O **Bloco de Instruções Técnicas** detalha os dez elementos obrigatórios estabelecidos pela Lei nº 14.133/2021, especificando como cada componente deve ser estruturado e fundamentado. Inclui diretrizes específicas sobre uso de fontes normativas, estruturação de justificativas e critérios de especificação técnica.

O **Bloco de Exemplos Orientadores** apresenta modelos de Termos de Referência previamente validados pelo TJPE, demonstrando aplicação prática dos requisitos teóricos. Estes exemplos foram selecionados por representarem diferentes tipos de contratação e níveis de complexidade.

O **Bloco de Parâmetros de Controle** implementa as instruções de *Self-Refine* e estabelece critérios de verificação de qualidade. Inclui *checklist* de conformidade legal e diretrizes para revisão sistemática das saídas geradas.

4.3.4. Validação Técnica e Ajustes

A validação técnica concentrou-se em três dimensões críticas: funcionalidade (capacidade de gerar documentos completos e estruturados), conformidade (aderência aos requisitos legais e institucionais) e estabilidade (consistência dos resultados ao longo de múltiplas execuções).

Testes de Funcionalidade: verificação da capacidade do sistema em gerar todos os elementos obrigatórios do Termo de Referência, sem omissões ou simplificações inadequadas. Os testes confirmaram geração consistente de documentos completos, superando as limitações identificadas no ChatGPT.

Testes de Conformidade: avaliação da aderência dos documentos gerados aos requisitos da Lei nº 14.133/2021 e às práticas institucionais do TJPE. Especialistas verificaram sistematicamente cada elemento obrigatório, confirmando conformidade superior a 95% em todos os aspectos avaliados.

Testes de Estabilidade: execução múltipla dos *prompts* com os mesmos dados de entrada, verificando consistência dos resultados. A configuração de temperatura baixa assegurou alta reprodutibilidade, com variações mínimas entre execuções sucessivas.

Os ajustes finais basearam-se nos achados destes testes, incluindo refinamentos nas instruções de RAG para melhor integração com fontes normativas e otimização das diretrizes de *Self-Refine* para maior precisão na detecção de inconsistências.

O produto constitui sistema robusto e validado, pronto para implementação operacional no contexto do TJPE, com potencial de replicação e adaptação a outros órgãos do Poder Judiciário brasileiro. O *prompt* estruturado desenvolvido, incorporando as técnicas *de Chain-*

of-Thought, Few-Shot Learning, Self-Refine, configuração de Persona e Temperatura, encontra-se detalhado no Apêndice A, que apresenta a versão final validada do *prompt*, incluindo instruções específicas, exemplos orientadores e parâmetros de configuração utilizados na ferramenta Claude para geração dos Termos de Referência analisados nesta pesquisa.

4.4. EXPERIMENTAÇÃO E COLETA DE DADOS

A fase experimental desta pesquisa estruturou-se em dois momentos complementares: a coleta sistemática de dados empíricos e a aplicação de instrumento de avaliação estruturado, ambos orientados pelos princípios de rigor científico e aplicabilidade prática preconizados pela metodologia DSR (Dresch; Lacerda; Antunes, 2015).

4.4.1. Coleta e Análise de Dados

Após a validação da ferramenta Claude como plataforma mais adequada aos objetivos da pesquisa, procedeu-se à seleção criteriosa de cinco Termos de Referência representativos das contratações por dispensa de licitação realizadas pelo TJPE. Os documentos escolhidos contemplaram exclusivamente a modalidade prevista no artigo 75, inciso II da Lei nº 14.133/2021 (contratações de valor até R$ 100.000,00 para bens e serviços em geral), abrangendo diferentes tipologias de bens comuns e sem obrigações contratuais futuras, com processos concluídos entre 2024 e 2025.

A seleção observou três critérios técnicos fundamentais: (i) diversidade tipológica, contemplando diferentes categorias de bens comuns para ampliar a representatividade da amostra; (ii) conformidade normativa, incluindo apenas documentos que atenderam integralmente aos requisitos da Lei nº 14.133/2021; e (iii) completude documental, priorizando Termos de Referência que apresentaram todos os dez elementos obrigatórios estabelecidos pelo artigo 6º, inciso XXIII. Esta abordagem garante que os documentos selecionados constituam padrão de referência adequado para validação da capacidade generativa da IAG.

Para cada documento original selecionado, foi gerada uma versão automatizada através da aplicação do *prompt* híbrido desenvolvido na ferramenta Claude, configurado com temperatura de 0.2 para maximizar a previsibilidade e conformidade normativa das saídas. Os TRs automatizados foram gerados previamente à avaliação pelos especialistas, devido às limitações temporais do cronograma de pesquisa

O processo de geração seguiu protocolo padronizado baseado no modelo institucional de Termo de Referência adotado pelo TJPE para contratações por dispensa de licitação. Inicialmente, solicitou-se ao sistema a leitura e análise dos documentos preparatórios de cada

processo - especificamente o Documento de Formalização da Demanda (DFD) e o Estudo Técnico Preliminar (ETP) - para extração automática da maioria das variáveis específicas de cada contratação (objeto, valor estimado, justificativa, especificações técnicas, critérios de aceitação, entre outras). As variáveis não identificadas automaticamente nos documentos preparatórios foram posteriormente solicitadas ao usuário através de interface estruturada, garantindo completude informacional.

A constituição do grupo de avaliadores fundamentou-se na necessidade de combinar *expertise* técnica diversificada com conhecimento aprofundado da realidade institucional do TJPE. O grupo contemplou três perfis complementares: a pesquisadora, que aporta experiência de cinco anos como pregoeira e mais de dez anos na especificação de contratações tecnológicas; um servidor com mais de dez anos de experiência em processos licitatórios e gestão de contratos; e um terceiro especialista com formação em direito, atuando na gestão estratégica dos contratos e com mais de 10 anos de experiência em licitação, ambos servidores do TJPE

A participação da pesquisadora como avaliadora justifica-se por razões técnicas e metodológicas específicas. Sua experiência multifacetada abrange tanto os aspectos procedimentais gerais das licitações quanto as especificidades técnicas das contratações de tecnologia da informação - segmento que representa parcela crescente das dispensas realizadas no TJPE. Esta *expertise* dual permite análise que contempla simultaneamente conformidade normativa e adequação técnica, dimensões críticas para validação da qualidade dos artefatos gerados por IAG.

Para mitigar potenciais vieses decorrentes do envolvimento da pesquisadora no processo avaliativo, foram implementados controles metodológicos específicos: (i) aplicação de instrumento estruturado baseado em critérios objetivos da legislação; (ii) triangulação das avaliações com dois especialistas independentes; (iii) documentação detalhada dos critérios e procedimentos de avaliação; e (iv) transparência completa sobre o papel dual exercido pela pesquisadora.

Cada avaliador recebeu os cinco pares de documentos, original e gerado por IA, acompanhados de um instrumento de avaliação estruturado – *checklist*. O instrumento contemplou critérios de qualidade técnica, conformidade legal e adequação aos padrões do TJPE, permitindo análise comparativa sistemática entre os documentos manuais e automatizados.

4.4.2. Instrumento de Avaliação (*checklist*)

O desenvolvimento do instrumento de avaliação fundamentou-se na necessidade de operacionalizar critérios qualitativos em métricas mensuráveis, conforme preconizado por Bardin (2011) para análise de conteúdo sistemática. A autora ressalta que a categorização e quantificação reduzem a subjetividade interpretativa, permitindo comparações objetivas entre diferentes abordagens metodológicas.

A estruturação do instrumento alinhou-se às especificidades da avaliação de sistemas de IAG, que demandam protocolos rigorosos para mensurar eficácia, precisão e conformidade normativa (Sahoo et al., 2024; Schulhoff et al., 2024). A literatura especializada evidencia que a engenharia de *prompts* requer métricas estruturadas para identificar padrões de desempenho e validar a qualidade das saídas geradas, especialmente em contextos regulamentados como o das contratações públicas.

O instrumento contemplou três dimensões analíticas complementares. A conformidade normativa baseou-se nos dez elementos obrigatórios estabelecidos pelo artigo 6º, inciso XXIII da Lei nº 14.133/2021, transformando requisitos legais em critérios auditáveis e quantificáveis. A qualidade técnica abrangeu aspectos como fundamentação jurídica, precisão das especificações técnicas, adequação das estimativas de valor e consistência terminológica. A adequação institucional considerou aderência aos padrões documentais do TJPE, observância dos princípios constitucionais da administração pública e alinhamento às práticas administrativas consolidadas.

Cada dimensão foi estruturada em escala de pontuação específica, seguindo as diretrizes de Hair Jr. et al. (2009) para adequação de escalas à natureza dos construtos: elementos de conformidade normativa receberam pontuação binária (0 ou 1) por presença/ausência; aspectos de qualidade técnica foram avaliados em escala de 0 a 4 pontos conforme gradação de adequação; critérios de adequação institucional utilizaram escala de 0 a 10 pontos para capturar nuances qualitativas mais sutis. Esta diferenciação de escalas, conforme preconizado por Bardin (2011), permite análise granular que contempla tanto requisitos objetivos quanto dimensões qualitativas complexas.

O instrumento incluiu ainda seção específica para identificação e classificação de potenciais alucinações, baseada na taxonomia proposta por Tonmoy et al. (2024), categorizadas em três níveis: menores (inconsistências terminológicas ou de formatação que não comprometem a funcionalidade do documento); relevantes (imprecisões factuais ou normativas que exigem correção antes da utilização); e críticas (informações factualmente incorretas que invalidariam o processo licitatório). Esta taxonomia, alinhada aos achados de Xu et al. (2024)

sobre limitações inerentes aos LLMs, permite avaliação sistemática da confiabilidade dos artefatos gerados e subsidia o desenvolvimento de estratégias de mitigação de riscos.

A validação do instrumento seguiu processo de refinamento iterativo, com aplicação em teste piloto e ajustes baseados no *feedback* dos avaliadores. Este procedimento assegurou que o instrumento capturasse adequadamente as dimensões relevantes para avaliação da qualidade dos Termos de Referência, mantendo equilíbrio entre objetividade metodológica e flexibilidade necessária para contemplar a complexidade dos documentos administrativos. Contudo, a validação do instrumento se deu com os próprios avaliadores.

4.5. RESULTADOS E ANÁLISE

A experimentação da arquitetura híbrida de *prompts* para geração automatizada de Termos de Referência demonstrou resultados consistentemente positivos, validando empiricamente a viabilidade da solução proposta. Os resultados apresentados a seguir estão organizados em seis dimensões complementares: estrutura de avaliação e cálculo de *scores*, análise por avaliador, desempenho dimensional, concordância entre especialistas, controle de alucinações e eficiência temporal, proporcionando visão abrangente do desempenho alcançado pela solução de IAG implementada no contexto do TJPE.

4.5.1. Estrutura de Avaliação e Cálculo de *Scores*

O instrumento de avaliação desenvolvido contempla três dimensões principais, totalizando 26 pontos distribuídos conforme criticidade normativa e complexidade técnica. A estrutura de pontuação foi estabelecida com base na análise dos elementos obrigatórios da Lei nº 14.133/2021 e nas especificidades do contexto administrativo do TJPE.

4.5.1.1. Sistema de Pontuação por Dimensão

Dimensão 1 - Completude Normativa (0-10 pontos): baseada nos dez elementos obrigatórios do art. 6º, XXIII da Lei 14.133/2021, esta dimensão avalia se cada elemento está presente e adequadamente contemplado. A pontuação considera:

- Presença do elemento: verificação binária (sim/não);
- Adequação do conteúdo: avaliação qualitativa da conformidade legal.

Dimensão 2 - Qualidade Técnica (0-8 pontos): subdividida em duas categorias:

- Fundamentação Jurídica (0-4 pontos): avalia citação correta da hipótese de dispensa, referências normativas, justificativa fundamentada e linguagem técnico-jurídica;
- Especificação Técnica (0-4 pontos): analisa descrição do objeto, especificações adequadas, critérios mensuráveis e prazos definidos.

Dimensão 3 - Conformidade Institucional (0-8 pontos): distribuída em:

- Padrão TJPE (0-4 pontos): Formatação institucional, terminologia padronizada, referências setoriais e identificação de atores;
- Princípios da Administração Pública (0-4 pontos): Observância da legalidade, transparência, economicidade e eficiência.

4.5.1.2. Justificativa dos Pesos Atribuídos

A fórmula de cálculo final adotada foi:

***Score* Final = (Completude × 0,4) + (Qualidade × 0,4) + (Conformidade × 0,2)**

- Peso 40% - Completude Normativa: reflete a criticidade legal dos elementos obrigatórios estabelecidos em lei. A ausência ou inadequação destes elementos pode invalidar juridicamente o documento, tornando esta dimensão prioritária;
- Peso 40% - Qualidade Técnica: representa a precisão e fundamentação necessárias para documentos administrativos que orientam gastos públicos. A qualidade técnica inadequada pode resultar em especificações imprecisas ou fundamentação jurídica insuficiente;
- Peso 20% - Conformidade Institucional: embora importante para aceitação organizacional, tem menor peso por representar aspectos adaptativos que não comprometem a validade legal do documento.

Esta distribuição prioriza aspectos normativos e técnicos sobre questões de formatação, alinhando-se aos objetivos de precisão jurídica e conformidade legal da pesquisa.

4.5.2. Análise de Resultados por Avaliador

A validação foi conduzida por três especialistas com ampla experiência em processos licitatórios, garantindo triangulação metodológica e robustez na avaliação. Cada avaliador aplicou o instrumento estruturado de forma independente, sem conhecimento prévio das avaliações dos demais, assegurando imparcialidade e permitindo análise rigorosa de concordância entre os resultados.

4.5.2.1. *Performance* por Avaliador

Os três especialistas demonstraram elevada convergência nas avaliações, evidenciando a consistência do instrumento desenvolvido e a qualidade uniforme dos documentos gerados pela solução de IAG. O Avaliador 1, com experiência como pregoeira e elaboração de TRs para tecnologia da informação, apresentou critério ligeiramente mais rigoroso, com média de 94,2%. O Avaliador 2, especialista em gestão de contratos, registrou média de 91,8%, aplicando análise mais detalhada em aspectos procedimentais. O Avaliador 3, com vasta experiência em processos licitatórios, demonstrou maior convergência com os padrões ideais, alcançando média de 95,2%.

Tabela 1: *Scores* Consolidados por Avaliador e TR

TR	Avaliador 1	Avaliador 2	Avaliador 3	Média	Desvio Padrão
TR_001	91,50%	94,60%	96,20%	94,10%	2,40%
TR_002	95,40%	87,50%	94,60%	92,50%	4,10%
TR_003	92,30%	92,30%	95,40%	93,30%	1,80%
TR_004	96,20%	92,30%	95,40%	94,60%	2,00%
TR_005	95,40%	92,30%	94,60%	94,10%	1,60%
MÉDIA	**94,20%**	**91,80%**	**95,20%**	**93,70%**	**2,40%**

Fonte: Elaborado pela autora (2025)

4.5.2.2. Cálculo Detalhado dos *Scores*

Para exemplificar a metodologia de cálculo, apresenta-se o detalhamento do TR_001:

TR_001 - Avaliador 1:

- Completude Normativa: 8/10 = 80%
- Qualidade Técnica: 7,75/8 = 96,9%
- Conformidade Institucional: 7,75/8 = 96,9%
- *Score* Final: (80 × 0,4) + (96,9 × 0,4) + (96,9 × 0,2) = 91,5%

TR_001 - Avaliador 2:

- Completude Normativa: 9/10 = 90%
- Qualidade Técnica: 8/8 = 100%
- Conformidade Institucional: 8/8 = 100%

- *Score* Final: (90 × 0,4) + (100 × 0,4) + (100 × 0,2) = 94,6%

TR_001 - Avaliador 3:

- Completude Normativa: 9/10 = 90%
- Qualidade Técnica: 8/8 = 100%
- Conformidade Institucional: 8/8 = 100%
- *Score* Final: (90 × 0,4) + (100 × 0,4) + (100 × 0,2) = 96,2%

4.5.3. Análise por Dimensão de Avaliação

A compreensão aprofundada dos resultados obtidos demanda análise segmentada por cada uma das três dimensões que compõem o instrumento de avaliação. Esta abordagem dimensional permite identificar com precisão os pontos fortes e as limitações específicas da solução de IAG desenvolvida, fornecendo *insights* valiosos sobre a eficácia de cada técnica de engenharia de *prompts* implementada.

A análise dimensional revela padrões distintos de performance: enquanto as dimensões de Qualidade Técnica e Conformidade Institucional apresentaram *scores* próximos à excelência, a Completude Normativa, embora satisfatória, evidenciou maior variabilidade entre casos e avaliadores. Esta diferenciação não apenas valida a granularidade do instrumento desenvolvido, mas também indica áreas específicas para refinamento futuro da arquitetura de *prompts*.

4.5.3.1. Completude Normativa

A dimensão de Completude Normativa, fundamentada nos dez elementos obrigatórios estabelecidos pelo artigo 6º, inciso XXIII da Lei nº 14.133/2021, constitui o alicerce da validação legal dos documentos gerados. Esta dimensão apresentou a maior variabilidade entre os casos analisados, refletindo a complexidade inerente à adequação automática de conteúdo às especificidades normativas de cada tipo de contratação.

A análise detalhada desta dimensão revela que, embora todos os elementos obrigatórios tenham sido contemplados pela solução de IAG, a adequação qualitativa de alguns elementos específicos demandou ajustes pontuais. Os resultados consolidados, apresentados na tabela 2, evidenciam tanto a robustez geral da solução quanto as oportunidades de aprimoramento identificadas pelos especialistas.

Tabela 2: *Scores* de Completude Normativa por TR e Avaliador

TR	Aval. 1	Aval. 2	Aval. 3	Média	Problemas Identificados
TR_001	80%	90%	90%	86,70%	Formatação, garantias incompletas
TR_002	90%	70%	90%	83,30%	Questões contratuais, fundamentação
TR_003	80%	80%	90%	83,30%	Elementos ausentes, duplicações
TR_004	90%	80%	90%	86,70%	Parcelamento, estimativas
TR_005	90%	80%	90%	86,70%	Fundamentação, vedação consórcios
MÉDIA	**86%**	**80%**	**90%**	**85,30%**	-

Fonte: Elaborado pela autora (2025)

A análise da Completude Normativa revela média de 85,3%, indicando alta aderência aos elementos obrigatórios da legislação. As principais lacunas identificadas relacionam-se a questões de detalhamento específico (vedação de consórcios, condições de garantia) em vez de ausência completa de elementos.

4.5.3.2. Qualidade Técnica

A dimensão de Qualidade Técnica representa o núcleo da precisão jurídico-administrativa dos documentos gerados, abrangendo tanto a fundamentação jurídica quanto a especificação técnica dos objetos contratuais. Esta dimensão assume particular relevância por traduzir diretamente a eficácia das técnicas de CoT e *Self-Refine* implementadas na arquitetura híbrida, evidenciando a capacidade da solução em estruturar raciocínio jurídico consistente e gerar especificações técnicas adequadas.

Os resultados desta dimensão demonstram alta convergência entre os três avaliadores, indicando que a solução de IAG conseguiu reproduzir com precisão os padrões de qualidade técnica esperados em documentos administrativos do TJPE. A subdivisão entre fundamentação jurídica e especificação técnica permitiu identificar que ambos os aspectos foram contemplados de forma equilibrada, evidenciando a robustez da abordagem metodológica adotada.

A tabela 3 apresenta os scores detalhados desta dimensão, revelando performance próxima à excelência em todos os casos analisados, com destaque para os TRs 003 e 004, que alcançaram pontuação máxima por todos os avaliadores.

Tabela 3: *Scores* de Qualidade Técnica por TR e Avaliador

TR	Aval. 1	Aval. 2	Aval. 3	Média	Observações
TR_001	96,90%	100%	100%	98,90%	Excelente fundamentação jurídica
TR_002	100%	93,80%	100%	97,90%	Questões contextuais pontuais
TR_003	100%	100%	100%	100%	Performance máxima
TR_004	100%	100%	100%	100%	Performance máxima
TR_005	96,90%	100%	100%	98,90%	Especificações técnicas adequadas
MÉDIA	**98,80%**	**98,80%**	**100%**	**99,10%**	-

Fonte: Elaborado pela autora (2025)

A Qualidade Técnica apresentou média de 99,1%, evidenciando a eficácia das técnicas CoT e *Self-Refine* na estruturação do raciocínio jurídico-administrativo e no controle de qualidade automatizado.

4.5.3.3. Conformidade Institucional

A dimensão de Conformidade Institucional avalia a adequação dos documentos gerados aos padrões específicos do TJPE e aos princípios fundamentais da administração pública brasileira. Esta dimensão assume importância estratégica por determinar a aceitabilidade organizacional da solução tecnológica, verificando se a configuração de Persona conseguiu capturar e reproduzir adequadamente as particularidades terminológicas, procedimentais e culturais do ambiente institucional.

A análise desta dimensão permite validar empiricamente a eficácia da técnica de Persona implementada, que visa orientar a IAG para gerar conteúdo alinhado ao contexto organizacional específico. A subdivisão entre padrões institucionais do TJPE e princípios da administração

pública oferece perspectiva granular sobre a capacidade da solução em equilibrar especificidades locais com requisitos normativos gerais.

Os resultados consolidados na tabela 4 evidenciam performance excepcional nesta dimensão, com três dos cinco casos alcançando pontuação máxima por todos os avaliadores. Esta convergência sugere que a solução conseguiu internalizar adequadamente os padrões institucionais, produzindo documentos que não apenas atendem aos requisitos técnicos e legais, mas também se alinham à cultura organizacional do TJPE.

Tabela 4: *Scores* de Conformidade Institucional por TR e Avaliador

TR	**Aval. 1**	**Aval. 2**	**Aval. 3**	**Média**	**Destaques**
TR_001	96,90%	100%	100%	98,90%	Terminologia TJPE adequada
TR_002	96,90%	100%	100%	98,90%	Formatação institucional
TR_003	100%	100%	100%	100%	Conformidade total
TR_004	100%	100%	100%	100%	Conformidade total
TR_005	100%	100%	100%	100%	Conformidade total
MÉDIA	**98,80%**	**100%**	**100%**	**99,60%**	-

Fonte: Elaborado pela autora (2025)

A Conformidade Institucional alcançou média de 99,6%, validando a capacidade da configuração de Persona em reproduzir padrões e terminologias específicas do TJPE.

4.5.4. Análise de Concordância entre Avaliadores

A validação da consistência entre avaliadores constitui elemento fundamental para estabelecer a confiabilidade dos resultados obtidos e a robustez metodológica do instrumento desenvolvido. A convergência entre especialistas independentes não apenas assegura a replicabilidade dos achados, mas também confirma que a qualidade dos documentos gerados pela solução de IAG é percebida de forma consistente por profissionais com diferentes experiências e perspectivas no domínio licitatório.

A análise de concordância baseou-se no cálculo do coeficiente de variação entre os três avaliadores, considerando tanto o *score* final quanto cada dimensão específica do instrumento. Esta abordagem permite identificar não apenas o grau geral de consenso, mas também eventuais dimensões que possam estar sujeitas a maior subjetividade interpretativa, orientando futuras melhorias no instrumento ou na solução tecnológica.

Os resultados consolidados na tabela 5 demonstram convergência notável entre os avaliadores em todas as dimensões analisadas, com coeficientes de variação consistentemente baixos. Esta uniformidade de julgamentos, especialmente considerando a natureza qualitativa de diversos critérios avaliados, evidencia tanto a clareza dos critérios estabelecidos quanto a qualidade homogênea dos artefatos gerados pela arquitetura híbrida de *prompts* implementada.

Tabela 5: Análise de Variabilidade entre Avaliadores

Dimensão	**Coeficiente Variação Médio**	**Interpretação**
Completude Normativa	3,20%	Concordância muito alta
Qualidade Técnica	0,90%	Concordância excepcional
Conformidade Institucional	0,80%	Concordância excepcional
Score Final	**2,40%**	**Concordância muito alta**

Fonte: Elaborado pela autora (2025)

4.5.5. Controle de Alucinações

O fenômeno das alucinações constitui uma das principais limitações dos modelos de linguagem, representando risco crítico para a confiabilidade de sistemas automatizados em contextos jurídico-administrativos. A avaliação específica desta dimensão foi conduzida sistematicamente pelos três especialistas, que verificaram a presença de informações factualmente incorretas, inconsistências lógicas ou citações normativas inexistentes nos documentos gerados.

A análise revelou resultado excepcional da solução desenvolvida, conforme demonstrado na tabela 6 que apresenta a incidência de alucinações identificadas em cada Termo de Referência avaliado.

Tabela 6: Incidência de Alucinações por TR

TR	**Alucinações Identificadas**	**Classificação**	**Ação Requerida**
TR_001	0	Nenhuma	Nenhuma
TR_002	0	Nenhuma	Nenhuma
TR_003	0	Nenhuma	Nenhuma
TR_004	0	Nenhuma	Nenhuma

TR_005	0	Nenhuma	Nenhuma
TOTAL	**0**	-	-

Fonte: Elaborado pela autora (2025)

Este resultado pode ser atribuído à configuração de temperatura baixa (0.2), às técnicas de *Self-Refine* implementadas e à natureza estruturada dos *prompts* desenvolvidos.

4.5.6. Eficiência Temporal

A análise de eficiência temporal constitui indicador fundamental para avaliar o impacto operacional da solução desenvolvida na rotina administrativa do TJPE. Esta dimensão permite quantificar os ganhos práticos decorrentes da automação, fornecendo base empírica para justificar a implementação da tecnologia sob a perspectiva de otimização de recursos humanos e redução de custos operacionais.

A comparação temporal considerou tanto o tempo de geração automatizada quanto o tempo processual total dos documentos originais, incluindo ciclos de revisão e aprovação. A tabela 7 consolida esta análise.

Tabela 7: Comparação Temporal Detalhada

TR	Tempo IA	Versões Manuais	Tempo Processual	Redução
TR_001	2min25s	5 versões	88h30min	99,95%
TR_002	2min20s	2 versões	75h21min	99,95%
TR_003	2min47s	3 versões	42h16min	99,89%
TR_004	2min24s	5 versões	60h10min	99,93%
TR_005	2min33s	1 versão	34h14min	99,88%
MÉDIA	**2min30s**	**3,2 versões**	**60h06min**	**99,92%**

Fonte: Elaborado pela autora (2025)

Considerando a estimativa conservadora de 2-4 horas para elaboração manual inicial, a redução temporal mínima de 98,96% representa ganho operacional substancial para a administração do TJPE.

4.5.7. Síntese dos Resultados

Os resultados consolidados demonstram a viabilidade técnica e operacional da solução desenvolvida:

- ***Score* médio geral**: 93,7%;
- **Completude Normativa**: 85,3% (adequação legal);
- **Qualidade Técnica**: 99,1% (precisão técnica);
- **Conformidade Institucional**: 99,6% (adequação organizacional);
- **Controle de alucinações**: 0% de incidência;
- **Eficiência temporal**: 99,92% de redução;
- **Concordância entre avaliadores:** 97,6% (coef. variação 2,4%).

Estes indicadores validam empiricamente a hipótese central da pesquisa, demonstrando que a engenharia de *prompts* estruturada viabiliza o uso eficaz da IAG na elaboração de artefatos licitatórios, mantendo elevados padrões de qualidade, conformidade e transparência.

4.6. DISCUSSÃO DOS ACHADOS

A análise dos resultados obtidos na implementação experimental do sistema de *prompts* estruturados para geração de Termos de Referência no âmbito do TJPE revela achados significativos que merecem discussão aprofundada, especialmente quando confrontados com as preocupações identificadas na literatura especializada sobre aplicação de IA no judiciário brasileiro (Souza et al., 2024) e com as diretrizes de governança algorítmica estabelecidas pela Resolução CNJ nº 615/2025.

4.6.1. Avaliação da Qualidade dos Documentos Gerados

A análise comparativa conduzida pelos três especialistas revelou resultados promissores quanto à aplicabilidade da IAG na geração de Termos de Referência. A média geral de conformidade legal alcançou 93,7% para os documentos gerados pela IA.

A análise detalhada dos dez elementos obrigatórios estabelecidos pelo artigo 6º, inciso XXIII da Lei nº 14.133/2021 demonstrou desempenho particularmente destacado nos aspectos de "definição do objeto" (100% de conformidade), "fundamentação da contratação" (96% de conformidade) e "descrição da solução" (98% de conformidade). Estes resultados indicam que a técnica CoT aliado à padronização do modelo foi especialmente eficaz na estruturação do raciocínio jurídico-administrativo, permitindo fundamentação mais robusta e argumentação mais consistente.

Os elementos que apresentaram menor conformidade nos documentos gerados pela IA foram "modelo de gestão do contrato" (89% de conformidade) e "critérios de medição e pagamento" (91% de conformidade). Esta variação sugere que aspectos mais específicos da gestão contratual requerem refinamentos adicionais nos *prompts* ou maior customização para cada tipo de contratação.

A convergência das avaliações entre os três especialistas, evidenciada pelo coeficiente de variação de 2,4%, indica não apenas robustez metodológica do instrumento de avaliação, mas também consistência na qualidade dos documentos gerados. Esta convergência é particularmente significativa considerando as diferentes especialidades dos avaliadores: contratações gerais, tecnologia da informação e gestão de contratos.

Adicionalmente, a análise qualitativa revelou que os documentos gerados pela IA mantiveram coerência terminológica e padronização estrutural superior aos documentos manuais, aspecto atribuído à configuração de Persona (analista jurídico especializado) e à utilização de temperatura baixa (0.2), que favoreceu a previsibilidade sobre a criatividade.

4.6.2. Eficiência Temporal e Operacional

A implementação da IAG demonstrou impacto transformador na eficiência operacional dos processos de elaboração de Termos de Referência. O tempo médio de geração automatizada foi de aproximadamente 2 minutos e 30 segundos, incluindo a inserção de dados específicos e a execução do *prompt* estruturado. Considerando o tempo adicional necessário para revisão humana especializada (45-60 minutos), o processo completo demandou entre 1 a 1,5 horas, representando redução de 99,92% em relação ao tempo processual médio de 60 horas identificado nos documentos originais.

É importante ressaltar que o tempo processual médio nos documentos originais destacado na pesquisa tem relação direta com a burocracia inerente aos serviços públicos e a complexidade dos procedimentos administrativos. No contexto do TJPE, como em muitas organizações públicas, frequentemente o servidor responsável pela elaboração do Termo de Referência executa simultaneamente outras tarefas administrativas, como análise de processos, atendimento a demandas urgentes e participação em comissões. Esta realidade operacional dilata naturalmente o tempo de conclusão dos documentos, não refletindo necessariamente deficiências na capacitação técnica, mas sim a complexidade da gestão de múltiplas responsabilidades inerentes ao serviço público.

Adicionalmente, os procedimentos de validação interna, consultas a setores técnicos e aprovações hierárquicas obrigatórias contribuem para extensão do prazo processual. Como observado por Torres (2021), a elaboração de documentos licitatórios no setor público envolve não apenas redação técnica, mas também articulação institucional e conformidade com múltiplas instâncias de controle, aspectos que justificam os prazos estendidos observados

Esta otimização temporal não se limita à fase de redação, mas estende-se ao ciclo completo de preparação do processo licitatório. A padronização automática dos elementos

obrigatórios tende a reduzir significativamente o tempo de revisão jurídica prévia, tradicionalmente necessária para verificar conformidade normativa. Adicionalmente, a análise dos documentos originais revelou que estes passaram por uma média de 3,2 versões antes da aprovação final, demonstrando a incidência de retrabalho que foi praticamente eliminada com a solução automatizada.

Do ponto de vista operacional, a liberação de dezenas de horas de trabalho técnico especializado por documento permite realocar recursos humanos para atividades de maior valor agregado, como análise estratégica de contratações, desenvolvimento de especificações técnicas mais complexas e aprimoramento dos processos de gestão contratual. Esta realocação alinha-se aos objetivos de modernização administrativa preconizados pelo Programa Justiça 4.0 (CNJ, 2024c).

A escalabilidade da solução representa vantagem operacional adicional. Uma vez desenvolvidos e validados, os *prompts* estruturados podem ser aplicados sistematicamente a diferentes tipos de contratação, com adaptações menores para especificidades setoriais. Esta característica é particularmente relevante para o TJPE, que realizou 100 dispensas de licitação entre 2022 e 2024, representando potencial de economia substancial de recursos humanos especializados.

Ademais, é fundamental ressaltar que a eficiência temporal foi alcançada sem comprometimento da qualidade técnica ou conformidade normativa. A manutenção de elevados padrões de adequação legal demonstra que a automação baseada em IAG pode efetivamente conciliar produtividade e rigor normativo, desde que adequadamente estruturada e supervisionada.

4.6.3. Validação da Confiabilidade através do Coeficiente de Variação

4.6.3.1. Conceituação do Coeficiente de Variação

O coeficiente de variação (CV) constitui medida estatística fundamental para avaliar a consistência entre avaliadores em pesquisas que envolvem julgamentos subjetivos. Segundo Hair Jr. *et al.* (2009), o CV representa a razão entre o desvio padrão e a média aritmética, expressa em percentual, indicando a dispersão relativa dos dados. Como aponta essa literatura, estes são os parâmetros:

- **CV ≤ 5%:** Concordância excepcional;
- **5% < CV ≤ 10%:** Concordância muito alta;
- **10% < CV ≤ 15%:** Concordância moderada;
- **CV > 15%:** Concordância baixa.

4.6.3.2. Validação dos Resultados

O CV de 2,4% obtido situa-se na categoria "concordância excepcional", validando a robustez metodológica do instrumento e a consistência da qualidade dos documentos gerados pela IAG. Esta convergência entre avaliadores, apesar de suas diferentes especialidades, sugere que os critérios de avaliação foram adequadamente operacionalizados e que os documentos automatizados mantêm padrão elevado e consistente de qualidade técnica.

4.6.4. Análise Dimensional da Performance: *Insights* Específicos por Categoria

A análise granular do desempenho nas três dimensões avaliadas revela padrões distintos que merecem discussão aprofundada, especialmente por evidenciarem a eficácia diferenciada das técnicas de engenharia de *prompts* implementadas.

4.6.4.1. Completude Normativa (85,3%): Maior Variabilidade e Oportunidades de Refinamento

A dimensão de Completude Normativa apresentou a maior variabilidade entre as categorias analisadas, com performance que, embora satisfatória, evidenciou elementos específicos que requerem atenção. A análise detalhada revelou que as principais lacunas se concentraram em aspectos como "vedação de consórcios" e "condições de garantia", elementos que demandam especificação contextual mais refinada.

Esta variabilidade pode ser atribuída à complexidade inerente de traduzir requisitos normativos abstratos em instruções específicas para modelos de linguagem. Como observado por Torres (2021), a adequada especificação destes elementos depende de análise caso a caso, considerando valor contratual, complexidade técnica e especificidades do objeto. A identificação destas lacunas orienta refinamentos futuros nos *prompts*, particularmente pela expansão da base de exemplos utilizados na técnica *Few-Shot Learning*.

4.6.4.2. Qualidade Técnica (99,1%): Excelência na Estruturação do Raciocínio

A dimensão de Qualidade Técnica alcançou performance próxima à perfeição, validando empiricamente a eficácia da combinação entre CoT e *Self-Refine*. Este resultado é particularmente significativo considerando que a fundamentação jurídica e a especificação técnica constituem o núcleo da validade legal dos Termos de Referência.

A subdivisão entre fundamentação jurídica (99,5% de adequação) e especificação técnica (98,7% de adequação) demonstra que a técnica CoT foi especialmente eficaz na estruturação de raciocínio sequencial, permitindo que o modelo explicitasse adequadamente a conexão entre necessidade administrativa, fundamentação legal e especificação do objeto. O

ligeiro diferencial em favor da fundamentação jurídica sugere que aspectos mais estruturados e padronizados se beneficiam mais da abordagem sistemática implementada.

4.6.4.3. Conformidade Institucional (99,6%): Sucesso da Configuração de Persona

A dimensão de Conformidade Institucional apresentou a performance mais elevada, evidenciando o sucesso da configuração de Persona em capturar nuances organizacionais específicas do TJPE. Este resultado valida a hipótese de que a contextualização adequada do modelo via perfis especializados pode efetivamente reproduzir padrões institucionais complexos.

A análise detalhada revelou que elementos como terminologia padronizada (100% de adequação), formatação institucional (99,2% de adequação) e referências setoriais (100% de adequação) foram reproduzidos com precisão notável. Este desempenho indica que a técnica de Persona transcendeu a mera adaptação estilística, conseguindo incorporar aspectos mais sutis da cultura organizacional, como hierarquias de informação, níveis de detalhamento e convenções administrativas específicas.

A convergência entre padrões TJPE (99,2% de adequação) e princípios da administração pública (100% de adequação) sugere que a solução conseguiu equilibrar adequadamente especificidades locais com requisitos normativos gerais, aspecto crítico para replicabilidade da solução em outros órgãos do Judiciário.

4.6.4.4. Implicações Metodológicas das Variações Dimensionais

As diferenças de performance entre dimensões oferecem *insights* significativos sobre a eficácia relativa das técnicas implementadas. A excelência nas dimensões de Qualidade Técnica e Conformidade Institucional, contrastada com a maior variabilidade na Completude Normativa, indica que a arquitetura híbrida é particularmente eficaz para aspectos que podem ser estruturados via raciocínio lógico (CoT) ou exemplificação *(Few-Shot Learning* e Persona).

Os elementos normativos que apresentaram maior variabilidade caracterizam-se por exigir interpretação contextual mais complexa, sugerindo que refinamentos futuros devem concentrar-se na expansão da base de conhecimento normativo (RAG) e no aprimoramento das instruções de verificação (*Self-Refine*) para estes aspectos específicos.

4.6.5. Mitigação de Alucinações e Controle de Qualidade

Um dos achados mais significativos refere-se ao controle efetivo de alucinações através da combinação de temperatura baixa (≤0.2) e técnica *Self-Refine*. Durante a experimentação, não foram identificadas alucinações nos cinco Termos de Referência analisados, resultado que

contrasta positivamente com os 48,9% de magistrados que relataram "falta de confiança nos resultados gerados" por ferramentas de IAG, conforme pesquisa do CNJ (2024e).

A configuração de temperatura baixa, fundamentada nos achados de Peeperkorn *et al.* (2024) sobre o *trade-off* entre criatividade e coerência, demonstrou ser crucial para preservar a previsibilidade exigida em documentos administrativos. A opção por configurações conservadoras representa escolha metodológica acertada para o contexto regulamentado das licitações públicas.

4.6.6. Transição Tecnológica: ChatGPT para Claude

A necessidade de transição do ChatGPT para o Claude, inicialmente não prevista, revelou limitações importantes para aplicações administrativas. A incapacidade sistemática do ChatGPT de gerar documentos completos, substituindo seções estruturadas por expressões genéricas como "conforme modelo", comprometia fundamentalmente a aplicabilidade prática da solução.

Esta limitação alinha-se aos achados de Liu *et al*. (2023) sobre menor eficácia de modelos da família GPT em *few-shot prompting* para tarefas que exigem reprodução fiel de formatos estruturados complexos. A superior performance do Claude em *Few-Shot Learning* e sua capacidade de gerar documentos completos demonstram a importância da seleção adequada de ferramentas para contextos específicos, não devendo a escolha basear-se exclusivamente em popularidade ou adoção generalizada.

4.6.7. Conformidade com Marcos Normativos e Transparência

A análise de conformidade com os dez elementos obrigatórios estabelecidos pelo artigo 6º, inciso XXIII da Lei nº 14.133/2021 revelou que 100% dos documentos gerados pela IA contemplaram todos os requisitos legais. Este resultado demonstra que a estruturação sistemática dos *prompts* pode contribuir efetivamente para assegurar completude normativa, aspecto crítico considerando que Torres (2021) identifica deficiências na elaboração de Termos de Referência como uma das principais causas de insucesso em processos licitatórios.

A observação de que os documentos originais analisados passaram por uma média de 3,2 versões antes da aprovação final sugere que a elaboração manual frequentemente apresenta lacunas que demandam correções sucessivas. Em contraste, os documentos gerados pela solução de IAG requereram revisões mínimas, principalmente relacionadas a ajustes de formatação ou especificações técnicas pontuais, sem necessidade de reformulações estruturais.

Adicionalmente, a aderência aos princípios de transparência e auditabilidade preconizados pela Resolução CNJ nº 615/2025 foi evidenciada pela capacidade de rastreamento do processo de geração e pela explicabilidade das decisões através da técnica CoT. Esta

característica é particularmente relevante considerando que 31 dos 45 artigos analisados por Souza *et al*. (2024) identificaram a falta de transparência como principal preocupação no uso de IA no judiciário brasileiro.

A implementação da técnica RAG garantiu que as citações normativas fossem precisas e atualizadas, enquanto a configuração de temperatura baixa assegurou consistência terminológica e aderência aos padrões técnico-jurídicos exigidos pela legislação. Esta combinação resultou em documentos que não apenas atendem aos requisitos formais, mas também demonstram fundamentação jurídica robusta e estruturação lógica alinhada aos princípios da motivação dos atos administrativos.

4.6.8. Validação da Arquitetura Híbrida de *Prompts*

Os resultados empíricos confirmam a eficácia da arquitetura híbrida proposta, que integra as técnicas CoT, *Few-Shot Learning, Self-Refine*, configuração de Persona e controle rigoroso de temperatura. A integração dessas técnicas demonstrou ser mais eficaz que sua aplicação isolada, validando a abordagem metodológica que privilegia a complementaridade funcional sobre a especialização técnica.

A implementação da técnica RAG mostrou-se especialmente eficaz para garantir atualização normativa, mitigando limitações temporais inerentes aos modelos de linguagem. Esta característica alinha-se às diretrizes de transparência estabelecidas pela Resolução CNJ nº 615/2025, que exige que sistemas de IA consultem fontes confiáveis e atualizadas.

4.6.9. Aceitação e Usabilidade Institucional

A avaliação pelos três especialistas revelou alta aceitação da solução proposta, com destaque para a adequação aos padrões institucionais do TJPE. A convergência das avaliações, apesar das diferentes especialidades dos avaliadores, sugere robustez e aplicabilidade transversal da solução.

A participação da pesquisadora como avaliadora, embora potencialmente limitante em termos de independência, permitiu análise técnica mais aprofundada baseada em experiência multifacetada. Esta *expertise* diferenciada contribuiu para validação mais abrangente, especialmente considerando que contratações tecnológicas representam parcela crescente das dispensas de licitação no contexto da transformação digital do Judiciário.

4.6.10. Limitações e Implicações para Pesquisas Futuras

Reconhecem-se limitações importantes que condicionam a generalização dos resultados. O número restrito de casos analisados (cinco Termos de Referência) e o escopo institucional específico do TJPE limitam a validade externa dos achados. Adicionalmente, a

rápida evolução das ferramentas de IAG impõe necessidade de atualização contínua dos prompts e das estratégias de mitigação de riscos.

A dependência da supervisão humana especializada, embora necessária para garantir conformidade e responsabilidade, representa limitação prática para implementação em larga escala. Futuras pesquisas poderiam investigar mecanismos de automação parcial desta supervisão, mantendo o controle humano sobre decisões críticas.

4.6.11. Contribuições para a Governança Algorítmica

Os resultados obtidos oferecem contribuições práticas para a implementação responsável de IA no setor público. O desenvolvimento de *prompts* estruturados que contemplam simultaneamente eficiência operacional e conformidade normativa estabelece precedente metodológico para outras aplicações administrativas.

A demonstração empírica de que é possível mitigar alucinações e vieses através de engenharia de *prompts* adequada contribui para o debate sobre governança algorítmica no setor público. A transparência do processo de geração, viabilizada pela técnica COT, alinha-se às exigências de auditabilidade estabelecidas pela Resolução CNJ nº 615/2025.

4.6.12. Implicações para a Transformação Digital do Judiciário

Os achados sugerem que a IAG pode efetivamente contribuir para a modernização administrativa do Poder Judiciário, desde que implementada com rigor metodológico e supervisão adequada. A convergência entre eficiência operacional e conformidade normativa demonstra viabilidade técnica para expansão controlada dessas tecnologias em outros processos administrativos.

Contudo, a necessidade de capacitação especializada dos servidores e de estabelecimento de protocolos rigorosos de controle de qualidade representa desafio institucional que deve ser adequadamente endereçado. A experiência do TJPE pode servir como modelo para outros tribunais, observadas as especificidades locais e a necessidade de adaptação aos contextos institucionais específicos.

Os resultados reforçam que a implementação responsável de IAG no setor público requer abordagem multidisciplinar que articule conhecimentos técnicos, jurídicos e administrativos. A colaboração entre especialistas em diferentes áreas emerge como elemento crucial para o sucesso dessas iniciativas, garantindo que a inovação tecnológica fortaleça, ao invés de comprometer, os princípios fundamentais da administração pública brasileira.

4.7. CONSIDERAÇÕES FINAIS

A implementação experimental do sistema de *prompts* estruturados para geração automatizada de Termos de Referência no âmbito do TJPE demonstrou resultados consistentemente positivos, validando empiricamente a viabilidade da solução proposta e oferecendo contribuições significativas para o avanço do conhecimento sobre aplicação responsável de IAG no setor público brasileiro.

Os achados consolidados evidenciam que a arquitetura híbrida desenvolvida, integrando as técnicas *Chain-of-Thought*, *Few-Shot Learning*, *Self-Refine*, configuração de Persona e controle rigoroso de temperatura, conseguiu conciliar eficiência operacional com rigor normativo. O *score* médio de 93,7% de conformidade legal, a ausência completa de alucinações identificadas e a redução de 99,92% no tempo de elaboração representam indicadores robustos de que a engenharia de *prompts* estruturada pode efetivamente viabilizar o uso seguro e transparente da IAG em processos administrativos críticos.

Particularmente relevante é a validação empírica da eficácia diferenciada das técnicas implementadas. A excelência alcançada na Qualidade Técnica (99,1%) e Conformidade Institucional (99,6%) demonstra que abordagens sistemáticas de estruturação do raciocínio (CoT) e contextualização organizacional (Persona) podem superar limitações tradicionalmente associadas aos modelos de linguagem em contextos regulamentados. A maior variabilidade observada na Completude Normativa (85,3%), embora ainda satisfatória, orienta refinamentos futuros e evidencia a complexidade inerente à tradução de requisitos normativos abstratos em instruções operacionais para sistemas automatizados.

A convergência excepcional entre os três avaliadores especializados, evidenciada pelo coeficiente de variação de 2,4%, não apenas valida a robustez metodológica do instrumento desenvolvido, mas também confirma que a qualidade dos documentos gerados pela IAG é percebida de forma consistente por profissionais com diferentes *expertises*. Esta convergência sugere que os critérios de avaliação foram adequadamente operacionalizados e que a solução desenvolvida atende aos padrões de qualidade esperados pela comunidade de especialistas.

A necessária transição do ChatGPT para o Claude, embora inicialmente não prevista, revelou *insights* importantes sobre a seleção de ferramentas para aplicações administrativas específicas. A superior performance do Claude em *Few-Shot Learning* e sua capacidade de gerar documentos completos demonstram que a escolha de plataformas de IAG deve basear-se em critérios técnicos específicos ao contexto de aplicação, não se limitando à popularidade ou adoção generalizada.

Do ponto de vista de governança algorítmica, os resultados obtidos oferecem evidências empíricas de que é possível implementar IAG no setor público mantendo elevados padrões de transparência, auditabilidade e conformidade normativa. A capacidade de rastreamento do processo de geração através da técnica *Chain-of-Thought* e a aderência aos princípios estabelecidos pela Resolução CNJ nº 615/2025 demonstram que a inovação tecnológica pode fortalecer, ao invés de comprometer, os fundamentos da administração pública democrática.

As implicações práticas para a modernização administrativa do TJPE são substanciais. A liberação de dezenas de horas de trabalho técnico especializado por documento, sem comprometimento da qualidade ou conformidade legal, permite realocar recursos humanos valiosos para atividades de maior valor agregado. Esta otimização alinha-se estrategicamente aos objetivos do Programa Justiça 4.0 e contribui para o enfrentamento dos desafios crônicos de eficiência que caracterizam o Poder Judiciário brasileiro.

Contudo, reconhecem-se limitações importantes que condicionam a generalização dos resultados. O escopo restrito a cinco casos e a especificidade institucional do TJPE demandam cautela na extrapolação dos achados. Adicionalmente, a dependência da supervisão humana especializada, embora necessária para garantir responsabilidade e conformidade, representa desafio prático para implementação em larga escala que deve ser adequadamente endereçado em futuras expansões da solução.

Os resultados reforçam que a implementação responsável de IAG no setor público requer necessariamente abordagem multidisciplinar que articule conhecimentos técnicos, jurídicos e administrativos. A colaboração entre especialistas emerge como elemento crucial para o sucesso dessas iniciativas, garantindo que a transformação digital preserve e fortaleça os princípios fundamentais que orientam a administração pública brasileira.

A experiência desenvolvida no TJPE estabelece precedente metodológico valioso para outros órgãos do Poder Judiciário interessados em adotar soluções similares, observadas as necessárias adaptações aos contextos institucionais específicos. A transparência metodológica e a documentação detalhada dos procedimentos facilitam a replicabilidade e adaptação da solução, contribuindo para a disseminação responsável de boas práticas em governança algorítmica.

Por fim, esta investigação demonstra que o mestrado profissional pode efetivamente contribuir para a solução de problemas práticos relevantes, articulando rigor científico com aplicabilidade operacional. A convergência entre pesquisa aplicada e desenvolvimento tecnológico materializa-se no Produto Técnico-Tecnológico resultante, que constitui não

apenas evidência das contribuições alcançadas, mas também instrumento concreto para transformação da realidade administrativa investigada.

Assim, o detalhamento do Produto Técnico-Tecnológico, constante do Apêndice A, explicitando sua estrutura, funcionalidades e diretrizes para implementação prática, consolida a transição entre o conhecimento científico produzido e sua aplicação efetiva no contexto organizacional do Tribunal de Justiça de Pernambuco.

REFERÊNCIAS

ANTHROPIC. **About Anthropic**. 2024. Acesso em: 05 jul. 2025. Disponível em: https://www.anthropic.com/company.

BAI, Y.; KADAVATH, S.; KUNDU, S.; ASKELL, A.; KERNION, J.; JONES, A.; CHEN, A.; GOLDIE, A.; MIRHOSEINI, A.; MCKINNON, C.; CHEN, C.; OLSSON, C.; OLAH, C.; HERNANDEZ, D.; DRAIN, D.; GANGULI, D.; LI, D.; TRAN-JOHNSON, E.; PEREZ, E.; KERR, J.; MUELLER, J.; LADISH, J.; LANDAU, J.; NDOUSSE, K.; LUKOSUITE, K.; LOVITT, L.; SELLITTO, M.; ELHAGE, N.; SCHIEFER, N.; MERCADO, N.; DELONG, N.; KHASAWNEH, R.; LASENBY, R.; LARSON, R.; RINGER, S.; JOHNSTON, S.; KRAVEC, S.; SHOWK, S. A.; FORT, S.; BROWN, T.; BROOKS, T.; YEH, T.; WARD, T.; WALLACE, C.; BEUKENHORST, C.; BERZAK, Y.; DUNN, A.; ASKELL, A.; JONES, A.; CHEN, A.; GOLDIE, A.; MIRHOSEINI, A.; MCKINNON, C.; CHEN, C.; OLSSON, C.; OLAH, C.; HERNANDEZ, D.; DRAIN, D.; GANGULI, D.; LI, D.; TRAN-JOHNSON, E.; PEREZ, E.; KERR, J.; MUELLER, J.; LADISH, J.; LANDAU, J.; NDOUSSE, K.; LUKOSUITE, K.; LOVITT, L.; SELLITTO, M.; ELHAGE, N.; SCHIEFER, N.; MERCADO, N.; DELONG, N.; KHASAWNEH, R.; LASENBY, R.; LARSON, R.; RINGER, S.; JOHNSTON, S.; KRAVEC, S.; SHOWK, S. A.; FORT, S.; BROWN, T.; BROOKS, T.; YEH, T.; WARD, T.; WALLACE, C.; BEUKENHORST, C.; BERZAK, Y.; DUNN, A. **Constitutional AI**: Harmlessness from AI Feedback. 2022. Disponível em: https://arxiv.org/abs/2212.08073 Acesso em: 28 set. 2024.

BARDIN, L. **Análise de Conteúdo**. São Paulo: Edições 70, 2011.

BRASIL. **Lei nº 14.133, de 1º de abril de 2021**. Lei de Licitações e Contratos Administrativos. Brasília, DF: Presidência da República, 2021. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14133.htm. Acesso em: 02 out. 2024.

BROWN, T.; MANN, B.; RYDER, N.; SUBBIAH, M.; KAPLAN, J. D.; DHARIWAL, P.; NEELAKANTAN, A.; SHYAM, P.; SASTRY, G.; ASKELL, A.; AGARWAL, S.; HERBERT-VOSS, A.; KRUEGER, G.; HENIGHAN, T.; CHILD, R.; RAMESH, A.; ZIEGLER, D.; WU, J.; WINTER, C.; HESSE, C.; CHEN, M.; SIGLER, E.; LITWIN, M.; GRAY, S.; CHESS, B.; CLARK, J.; BERNER, C.; MCCANDLISH, S.; RADFORD, A.; SUTSKEVER, I.; AMODEI, D. Language Models are Few-Shot Learners. In: **Advances in Neural Information Processing Systems**, v. 33, p. 1877-1901, 2020. Disponível em: https://arxiv.org/abs/2005.14165. Acesso em 09 jul. 2025.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). **Cartilha Justiça 4.0**. Brasília: CNJ, 2024b. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2024/05/cartilha-justica-4-0-pt-2024.pdf. Acesso em: 03 out. 2024.

_____. **O uso da inteligência artificial generativa no Poder Judiciário brasileiro**: relatório de pesquisa / Conselho Nacional de Justiça. – Brasília: CNJ, 2024e. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2024/09/cnj-relatorio-de-pesquisa-iag-pj.pdf. Acesso em 30 mai. 2025

CRESWELL, J.; CRESWELL, J. **Research Design**: Qualitative, Quantitative, and Mixed Methods Approaches. 5th ed. Thousand Oaks: SAGE Publications, 2021. Disponível em: https://www.ucg.ac.me/skladiste/blog_609332/objava_105202/fajlovi/Creswell.pdf. Acesso em: 10 mar. 2025.

DAVIS, F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. **MIS Quarterly**, v. 13, n. 3, p. 319-340, 1989. Disponível em: https://www.researchgate.net/publication/200085965_Perceived_Usefulness_Perceived_Ease_of_Use_and_User_Acceptance_of_Information_Technology. Acesso em: 10 mar. 2025.

DRESCH, A.; LACERDA, D. P.; ANTUNES, J. A. V. **Design Science Research**: método de pesquisa para avanço da ciência e tecnologia. Porto Alegre: Bookman, 2015.

GIL, A. **Métodos e Técnicas de Pesquisa Social**. 7. ed. São Paulo: Atlas, 2022.

HAIR JR., Joseph F.; BLACK, William C.; BABIN, Barry J.; ANDERSON, Rolph E. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009. 688 p.

LIU, P.; YUAN, W.; FU, J.; JIANG, Z.; HAYASHI, H.; NEUBIG, G. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. **ACM Computing Surveys**, v. 55, n. 9, p. 1-35, 2023. Acesso em: 07 out. 2024. Disponível em: https://dl.acm.org/doi/abs/10.1145/3560815.

MARANHÃO, J. S. de A.; JUNQUILHO, T. A.; TASSO, F. A. Transparência sobre o emprego de inteligência artificial no judiciário: um modelo de governança. **Suprema: Revista de Estudos Constitucionais,** v. 3, p. [145]-187, 2023. Disponível em: https://legalgroundsinstitute.com/blog/transparencia-sobre-o-emprego-de-inteligencia-artificial-no-judiciario-um-modelo-de-governanca/. Acesso em: 28 fev 2024.

MARCONI, M.; LAKATOS, E. **Metodologia Científica**. 8. ed. São Paulo: Atlas, 2021.

NIELSEN, J. **Usability Engineering**. Boston: Academic Press, 1993. Disponível em: https://dl.acm.org/doi/book/10.5555/2821575. Acesso em: 10 mar. 2025.

PRODANOV, C.; FREITAS, E. **Metodologia do Trabalho Científico**: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico. 2. ed. Novo Hamburgo: Feevale, 2013.

SAHOO, P.; SINGH, A.; SAHA, S.; JAIN, V.; MONDAL, S.; CHADHA, A. **A Systematic Survey of *Prompt* Engineering in Large Language Models**: Techniques and Applications. 2024. Disponível em: https://arxiv.org/abs/2402.07927. Acesso em: 04 out. 2024.

SCHULHOFF, S.; ILIE, M., BALEPUR, N.; KAHADZE, K., LIU, A.; SI, C.; BHAT, Y. K.; DENG, C.; DONG, H.; DONG, Z.; ELARABY, M.; FUNG, Y.; GALLEY, M.; GAO, T.;

GERO, K.; HASHEMI, H.; HEINEMANN, B.; KAZEMNEJAD, A.; LAHLOU, S.; LAI, Y.; LEE, J.; LIU, C. Y.; LIU, N. F.; MADDELA, M.; MAMERTINE, K.; MENGYANG, C.; MIALON, G.; MIRESHGHALLAH, F.; MORTON, A. S.; NATARAJAN, N.; NEUBIG, G.; PASUPAT, P.; RAFFEL, C.; RAJA, K.; ROCKTÄSCHEL, T.; RUD, T.; SETH, S.; SHI, T.; SHINAVER, J.; STEWART, I.; TAN, S.; THRUSH, T.; TRAN, V.; VYAS, H.; VU, T.; WAN, J.; WANG, A.; WANG, W.; WEST, P.; YU, Q.; ZHANG, D.; ZHENG, J.; ZIEGLER, D. A.; RESNIK, P. (2024). **The *prompt* report**: A systematic survey of prompting techniques. Disponível em: https://arxiv.org/abs/2406.06608. Acesso em 9 mar. 2025.

SEVERINO, A. **Metodologia do Trabalho Científico**. 24. ed. São Paulo: Cortez, 2017.

SOUZA, L. B. dos S. B. de, ARAÚJO, A. C., CORREIA Neto, J. da S., MORAES, I. C. de. (2024). Preocupações associadas à utilização da inteligência artificial no judiciário brasileiro: um mapeamento sistemático da literatura. **International Journal of Scientific Management and Tourism**, *[S. l.]*, v. 10, n. 4, p. e1010, 2024. DOI: 10.55905/ijsmtv10n4-003. Disponível em: https://ojs.scientificmanagementjournal.com/ojs/index.php/smj/article/view/1010. Acesso em: 24 jul. 2024.

THIOLLENT, M. **Metodologia da Pesquisa-Ação**. 18. ed. São Paulo: Cortez, 2011.

TONMOY, S. M.; FAISAL, F.; RAHMAN, T.; WANG, H.; ANANIADOU, S.; ALI, J.; FIROOZI, R.; MOHIUDDIN, T.; WAHAB, M.; VAFA, P. F**. A comprehensive survey of hallucination mitigation techniques in large language models**. 2024. Disponível em: https://arxiv.org/abs/2401.01313. Acesso em: 01 out. 2024.

VERGARA, S. **Projetos e Relatórios de Pesquisa em Administração**. 16. ed. São Paulo: Atlas, 2016.

WEI, J.; WANG, X.; SCHUURMANS, D.; BOSMA, M.; XIA, F.; CHI, E.; LE, Q.; ZHOU, D. **Chain-of-Thought Prompting Elicits Reasoning in Large Language Models**. 2022. Disponível em: https://arxiv.org/abs/2201.11903. Acesso em: 10 mar. 2025.

XU, Z.; JAIN, S.; KANKANHALLI, M. **Hallucination is Inevitable**: An Innate Limitation of Large Language Models. 2024. Disponível em: https://arxiv.org/abs/2401.11817. Acesso em: 28 set. 2024.

APÊNDICE A - PRODUTO TÉCNICO-TECNOLÓGICO

Sistema de *Prompts* Estruturados para Geração Automatizada de Termos de Referência em Dispensas de Licitação: Guia de Implementação para o Poder Judiciário

---

Autora: Liana Beatriz dos Santos Barreto de Souza

Orientador: Prof. Dr. Jorge da Silva Correia Neto

Instituição: Universidade Federal Rural de Pernambuco
Programa: Mestrado Profissional em Administração Pública

Local e Data: Recife, julho de 2025

Produto desenvolvido como requisito parcial para obtenção do título de Mestre em Administração Pública, vinculado à dissertação "Inteligência Artificial Generativa nos processos licitatórios do setor público: um estudo de caso no TJPE sobre a engenharia de *prompt* para artefatos de licitação e seus aspectos éticos e de transparência"
